

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.  
OM protein - protein search, using sw model  
Run on: September 23, 2004, 23:33:46 ; Search time 73.5819 Seconds  
(without alignments)  
1509.371 Million cell updates/sec

Title: US-09-913-767-9  
Perfect score: 1760  
Sequence: 1 MKALVILNCIIAIGNCGG.....QBESEOTQSSLSRPISEC 352

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues  
Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 200000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPTREMBL\_25:\*

- 1: sp\_archaea:\*
- 2: sp\_bacteria:\*
- 3: sp\_fungi:\*
- 4: sp\_human:\*
- 5: sp\_invertebrate:\*
- 6: sp\_mammal:\*
- 7: sp\_mhc:\*
- 8: sp\_organelle:\*
- 9: sp\_phase:\*
- 10: sp\_plant:\*
- 11: sp\_rodent:\*
- 12: sp\_virus:\*
- 13: sp\_vertebrate:\*
- 14: sp\_unclassified:\*
- 15: sp\_rvirus:\*
- 16: sp\_bacteriap:\*
- 17: sp\_archaeap:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES				
Result No.	Score	Query Match %	Length	Description
1	1744.5	99.1	351	Q9fz95 arabidopsis
2	1149	65.3	356	P93010 arabidopsis
3	1149	65.3	358	Q94gb1 arabidopsis
4	1128	64.1	356	Q9fz96 arabidopsis
5	1124	63.9	356	Q9m7r1 arabidopsis
6	691.5	39.3	399	Q8h7p2 oryza sativ
7	557.5	31.7	344	O49726 arabidopsis
8	557	31.6	377	Q8ry74 arabidopsis
9	557	31.6	1128	O49724 arabidopsis
10	556.5	31.6	379	Q9lpf6 arabidopsis
11	555.5	31.6	348	O49725 arabidopsis
12	534	30.3	751	Q7x7p4 oryza sativ
13	530	30.1	315	Q9zuh3 arabidopsis
14	516.5	29.3	382	Q9sy29 arabidopsis
15	502.5	28.6	358	O49722 arabidopsis
16	492	28.0	372	Q8s170 oryza sativ

17	475	27.0	361	10	Q8RY83	Q8RY83 arabidopsis
18	475	27.0	432	10	Q9LDT1	Q9LDT1 arabidopsis
19	472.5	26.8	384	10	Q7XN26	Q7XN26 oryza sativ
20	458.5	26.1	358	10	Q9FLA4	Q9FLA4 arabidopsis
21	357	20.3	393	10	Q9FXH5	Q9FXH5 arabidopsis
22	323	18.4	232	10	Q8H7F0	Q8H7F0 arabidopsis
23	316	18.0	394	10	Q9C654	Q9C654 arabidopsis
24	289	16.4	387	10	Q9LQZ0	Q9LQZ0 arabidopsis
25	282	16.0	390	10	Q9C508	Q9C508 arabidopsis
26	279.5	15.9	392	10	Q9SX93	Q9SX93 arabidopsis
27	277.5	15.8	389	10	Q9SX90	Q9SX90 arabidopsis
28	262.5	14.9	398	10	Q9C655	Q9C655 arabidopsis
29	261.5	14.9	383	10	O04508	O04508 arabidopsis
30	229.5	13.0	149	10	Q94JW7	Q94JW7 arabidopsis
31	156	8.9	299	2	Q8KH10	Q8KH10 bacillus th
32	149	8.5	317	10	Q7XVK2	Q7XVK2 oryza sativ
33	143.5	8.2	335	10	Q9LIR9	Q9LIR9 arabidopsis
34	141.5	8.0	328	10	Q8GWX2	Q8GWX2 arabidopsis
35	140.5	8.0	323	10	Q9SKQ1	Q9SKQ1 arabidopsis
36	139.5	7.9	364	5	Q8IN50	Q8IN50 drosophila
37	139.5	7.9	578	5	Q8TOM8	Q8TOM8 drosophila
38	137.5	7.8	333	10	Q8LA45	Q8LA45 arabidopsis
39	136	7.7	368	10	Q9LWK7	Q9LWK7 arabidopsis
40	135	7.7	435	16	Q87GA5	Q87GA5 vibrio para
41	134.5	7.6	368	10	Q9SUFO	Q9SUFO arabidopsis
42	134.5	7.6	477	16	Q81XE8	Q81XE8 bacillus an
43	134	7.6	461	16	Q97E21	Q97E21 clostridium
44	133.5	7.6	345	16	Q8Y777	Q8Y777 listeria mo
45	133	7.6	294	16	Q81MZ5	Q81MZ5 bacillus an

ALIGNMENTS

RESULT 1

ID	Q9FZ95	PRELIMINARY;	PRT;	351 AA.
AC	Q9FZ95;			
DT	01-MAR-2001	(TEMBLrel. 16, Created)		
DT	01-MAR-2001	(TEMBLrel. 16, Last sequence update)		
DT	01-JUN-2003	(TEMBLrel. 24, Last annotation update)		
DE	F3H9.12	protein.		
GN	F3H9.12	Arabidopsis thaliana (Mouse-ear cross).		
OS	Arabidopsis thaliana (Mouse-ear cross).			
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;			
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;			
OC	eurosidb II; Brassicales; Brassicaceae; Arabidopsi			
OX	NCBI_TaxID=3702;			
RN	[1]			
RP	SEQUENCE FROM N.A.			
RA	Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,			
RA	Altai H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,			
RA	Buehler E., Chao Q., Chin C., Chou J., Choi E., Gonzalez A.,			
RA	Hwang B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,			
RA	Lenz C., Liu A., Liu S., Mukharsy N., Pham P., Sakano H., Shinn P.,			
RA	Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;			
RL	Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.			
DR	EMBL; AC021044; AAF98433.1; -			
DR	PIR; C86408; C86408.			
DR	InterPro; IPR004853; DUF250.			
DR	Pfam; PF03151; DUF250; 1.			
SQ	SEQUENCE 351 AA; 38871 MW; 992F1704CB76ED70 CRC64;			
Query Match 99.1%; Score 1744.5; DB 10; Length 351;				
Best Local Similarity 99.7%; Pred. No. 5.1e-118;				
Matches 351; Conservative 0; Mismatches 0; Indels 1; Gaps 1;				
QY	1	MKALVILNCIIAIGNCGGLIMRLYFNNNGKRIWFSTFLETAGFPVIFLPFLFYIR	60	
Db	1	MKALVILNCIIAIGNCGGLIMRLYFNNNGKRIWFSTFLETAGFPVIFLPFLFYIR	60	
QY	61	RRSNVGDSTSFPLIKPRLLIAAVIVGILSGDNVLYAYGIAYLPVSTAAALIASQLAFI	120	

	Query Match	Best Local Similarity	65.3%;	Score 1149;	DB 10;	Length 356;
	Matches 218;	Conservative	60.4%;	Pred. No. 4.3e-75;	Mismatches 57;	Indels 18;
						Gaps 3;
Qy	1	MVKALVIINCIILAIAGNCGGPIIMRLYFNNGKRIWFSTFLTAGTGPVIFPLLESYITR	60			
Db	1	MKTVLVINCIIFLAIGNCGGPLMMRLYFQNGGERIWFPSFLQTGVCPLITFFPLLSFLRR	60			
Qy	61	RRSNNVGDSTSFLLIKPRLLIAAVVGIISGFDNYLYAYGIAIYLVYSTAAIIIASQLAFI	120			
Db	61	RRCLREQETTFPFLMKPPLFIAIIVVGLLVGFDNYLYSYGLAYIPVSTASLLIISQLGFT	120			
Qy	121	AIFSFFWKVKHPTFFFINAVVLTVGCAVLGMHETETDKPVHETHKQIYITGFLITVAAMVM	180			
Db	121	ALFAFMVKQKTFPTFINAIVLTVGAVVLAMNSDSDLKANETHKEYVYVVGFTMLGALL	180			

```

181 YAFILPLVELAYOKAKOTMSYTLVLEFQLTCLLASIVSGVMFIAGDFK----- 230
181 YGFIPLPELVLSYKKSQRITYYTLLALEFQWVLCFAATCVCLVGLAAGDFKVGHALFIFKN 240
231 QALPKREARFKLGEALFYVAVFSAIIWQFFLGAIGLIFSTSLVSGIMISVLLPITEV 290
241 RVIAAGEARDPKLGSLEYVYVIVFTAIWQAFFVGAIGLIFCASLSVSGIMVWALLPVTVI 300
291 LAVIFYHEKFOAREKGLSLALSLSWGSFYFYGEIKSGEDKRRIOQEESEQTEOSSLSRPIS 350
301 LAVICFOEKFQAKGVALALSLSWGSVSYFYGVQKSEE-----KTKAQDTQLSQL--PVT 352
351 E 351
353 D 353

RESULT 3
Q94GB1 PRELIMINARY; PRT; 358 AA.
ID Q94GB1
AC Q94GB1, 2001 (Tremblrel. 19, Created)
DT 01-DEC-2001 (Tremblrel. 12, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE Putative purine permease (Fragment).
GN PUP2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatopsida; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=CV. C 24;
RX MEDLINE=20129770; PubMed=10662864;
RA Gillissen B., Burkle L., Andre B., Kuhn C., Rentsch D., Brandl B.,
RA Frommer W.B.;
RT "A new family of high-affinity transporters for adenine, cytosine, and
RT purine derivatives in Arabidopsis.";
RL Plant Cell 12:291-300(2000).
DR EMBL; AF078532; AAK61813.1; -.
DR InterPro; IPR004853; DUF250.
DR Pfam; PF03151; DUF250; 1.
DR NON TER 358
FT SEQUENCE 358 AA; 39461 MW; 69709A0EA6ECD6A8 CRC64;
SQ
Query Match 65.3%; Score 1149; DB 10; Length 358;
Best Local Similarity 60.4%; Pred. No. 4.3e-75;
Matches 218; Conservative 68; Mismatches 57; Indels 18; Gaps 3
QY 1 MVKALVIINCIILAINGCGGLIMRIYFNNGKRIWFSTFLETAGPPIFILLFSVITR 60
Db 3 MKTVLVIINCIIFLAINGCGGLMRLYFQNGGRIWFPSPQLQVGCPLIFPULLSFLRR 62
QY 61 RRSNNVGDSTSFELIKPRLLIAAVVIGILSGFDNYLYAYGIALYVSTAAIIASQAFI 120
Db 63 RCRLCEQETTPFLMKPPLFIAIVVGLLVGFNDNYLSYGLAIVPSTASLIISAQLGFT 122
QY 121 AIFSPFWVKHKTPPTFINAVVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVAAMV 180
Db 123 ALFAFWVKQKTPPTFINAVVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVAAMV 182
QY 181 YAFILPLVELAYOKAKOTMSYTLVLEFQLTCLLASIVSGVMFIAGDFK----- 230
Db 183 YGFIPLPELVLSYKKSQRITYYTLLALEFQWVLCFAATCVCLVGLAAGDFKVGHALFIFKN 240
QY 231 QALPKREARFKLGEALFYVAVFSAIIWQFFLGAIGLIFSTSLVSGIMISVLLPITEV 290
Db 243 RVIAAGEARDPKLGSLEYVYVIVFTAIWQAFFVGAIGLIFCASLSVSGIMVWALLPVTVI 300
QY 291 LAVIFYHEKFOAREKGLSLALSLSWGSFYFYGEIKSGEDKRRIOQEESEQTEOSSLSRPIS 350
Db 301 LAVICFOEKFQAKGVALALSLSWGSVSYFYGVQKSEE-----KTKAQDTQLSQL--PVT 352
QY 351 E 351
Db 353 D 353

```

```

Db 303 LAVICPOEKEQAGKGVVALSLWGSVSYFYGVKSEB-----KTKAQDTQLSQL--PVT 354
QY 351 E 351
Db 355 D 355

RESULT 4
Q9FZ96 PRELIMINARY; PRT; 356 AA.
ID Q9FZ96
AC Q9FZ96;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Purine permease.
GN F3H9.22 OR AT1G28230/F3H9 10 OR AT1G28230.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Pederspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.P.,
RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler B., Chao Q., Chin C., Chlou J., Choi E., Gonzalez A.,
RA Howng B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharkey N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Seki M., Iida K., Satou M., Sakurai T., Akiyama K., Ishida J.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RT "Arabidopsis thaliana full-length cDNA.";
RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
RN [3]
SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Juan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC021044; AAF98432.1; -.
DR EMBL; AK117664; BAC42317.1; -.
DR EMBL; BT005504; AAO63924.1; -.
DR PIR; D86408; D86408.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR004853; DUF250.
DR InterPro; IPR006020; DUF6.
DR Pfam; PF03151; DUF250; 1.
DR Pfam; PF00892; DUF6; 1.
SQ SEQUENCE 356 AA; 35280 MW; FA2701E954CF797B CRC64;

Query Match 64.1%; Score 1128; DB 10; Length 356;
Best Local Similarity 60.1%; Pred. No. 1.4e-73;
Matches 215; Conservative 66; Mismatches 65; Indels 12; Gaps 5;

QY 1 MKVALVIINCIILAIAGCGPLIMRLYFNNGKRIWFSTFLETAGPVPVIFIPLLFSYIR 60
Db 1 MKNGLIIINCIILITGTCGGLLRLYFTNGGKRIWFMSFLSTAGPFIILPLVLSFLSR 60
QY 61 RRSNNVGDG-----TSFELIKPRLLIAAVVIGLSGFDNLYVAYGIAYLPVSTAALIA 114
Db 61 RGRNRPNNAENKRTKLFMETPLFIASIVIGLLTGLDNLVSYGLAYLPVSTSSLIIG 120
QY 115 SQLAFIAIFSPFMVGHKFTPTTINAVLLTVGAALVGMHTETDKPVHETHKQVITGLIT 174
Db 121 TQLAFNALFAFLVVKQKTPPFSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
QY 175 VAAAVMYATILPLVELAYOKAKQTMSTYLVLEFQLILCLLASIVSVIGMEIADGFKQALP 234
Db 181 VVAALYAFILPLVELTYKKARQETFFPLVLEIQMWCLAAATFFCVIGMFVGDGFK-VIA 239
QY 235 KEAREFKLGEALF-YVAVFSAIIWQGFGLGALIFSTSSLSVSGIMISVLLPITVLA 293
Db 240 REAREFKIGSVFYFALIVITGIWQGFGLGALIVFVCASSLASGVLLSVLPVTEFAV 299
QY 294 IFYHEKFOAEKGLSLALSLSLWGVSYFYGEIKSGEDKRRIQOEBESQTEOSSLSRPISE 351
Db 300 VCFREKFOAEKGLSLALSLSLWGVSYFYGEIKSG--KKVVDKQPQETELPIL--PVSD 353

RESULT 5
Q9M7R1 PRELIMINARY; PRT; 356 AA.
ID Q9M7R1
AC Q9M7R1;
DT 01-OCT-2000 (T-EMBLrel. 15, Created)
DT 01-OCT-2000 (T-EMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Purine permease.
GN PUP1.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA MEDLINE=20129770; PubMed=10652864;
RA Gillissen B., Burkle L., Andre B., Kuhn C., Rentsch D., Brandl B.,
RA Frommer W.B.;
RT "A new family of high-affinity transporters for adenine, cytosine, and
RT purine derivatives in Arabidopsis.";
RL Plant Cell 12:291-300(2000).
DR EMBL; AF078531; AAF64547.1; -.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR004853; DUF250.
DR InterPro; IPR006020; DUF6.
DR Pfam; PF03151; DUF250; 1.
DR Pfam; PF00892; DUF6; 1.
SQ SEQUENCE 356 AA; 39250 MW; 0A38267726081261 CRC64;

Query Match 63.9%; Score 1124; DB 10; Length 356;
Best Local Similarity 59.8%; Pred. No. 2.7e-73;
Matches 214; Conservative 66; Mismatches 66; Indels 12; Gaps 5;

QY 1 MKVALVIINCIILAIAGCGPLIMRLYFNNGKRIWFSTFLETAGPVPVIFIPLLFSYIR 60
Db 1 MKNGLIIINCIILITGTCGGLLRLYFTNGGKRIWFMSFLSTAGPFIILPLVLSFLSR 60
QY 61 RRSNNVGDG-----TSFELIKPRLLIAAVVIGLSGFDNLYVAYGIAYLPVSTAALIA 114
Db 61 RGRNRPNNAENKRTKLFMETPLFIASIVIGLLTGLDNLVSYGLAYLPVSTSSLIIG 120
QY 115 SQLAFIAIFSPFMVGHKFTPTTINAVLLTVGAALVGMHTETDKPVHETHKQVITGLIT 174
Db 121 TQLAFNALFAFLVVKQKTPPFSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
QY 175 VAAAVMYATILPLVELAYOKAKQTMSTYLVLEFQLILCLLASIVSVIGMEIADGFKQALP 234
Db 181 VVAALYAFILPLVELTYKKARQETFFPLVLEIQMWCLAAATFFCVIGMFVGDGFK-VIA 239
QY 235 KEAREFKLGEALF-YVAVFSAIIWQGFGLGALIFSTSSLSVSGIMISVLLPITVLA 293
Db 240 REAREFKIGSVFYFALIVITGIWQGFGLGALIVFVCASSLASGVLLSVLPVTEFAV 299
QY 294 IFYHEKFOAEKGLSLALSLSLWGVSYFYGEIKSGEDKRRIQOEBESQTEOSSLSRPISE 351
Db 300 VCFREKFOAEKGLSLALSLSLWGVSYFYGEIKSG--KKVVDKQPQETELPIL--PVSD 353

```

300 VCFREKFOAEKGVSLLSLWGFVSFYGEFKSG--KKVVDKPPQPTETELPIL--PVSD 353

RESULT 6

Q8H7P2 PRELIMINARY; PRT; 399 AA.  
 AC Q8H7P2; (TREMRELrel. 23, Created)  
 DT 01-JUN-2003 (TREMRELrel. 23, Last sequence update)  
 DT 01-OCT-2003 (TREMRELrel. 24, Last annotation update)  
 DE Hypothetical protein.  
 GN Q1217B09.4.  
 OS Oryza sativa (japonica cultivar-group).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
 OC Ehrhartoideae; Oryzaceae; Oryza.  
 OX NCBI\_TaxID=39947;  
 [1]  
 RP SEQUENCE FROM N.A.  
 RA Wing R.A., Yu Y., Soderlund C., Kim H.-R., Rambo T., Saski C.,  
 RA Currie J., Collura K.; "  
 RT "Rice Genomic Sequence.";  
 RL Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AC121489; AAN64138.1; -;  
 DR InterPro; IPR004853; DUF250.  
 DR Pfam; PF03151; DUF250; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 399 AA; 41487 MW; D180480A4081416A CRC64;

Query Match 39.3%; Score 691.5; DB 10; Length 399;  
 Best Local Similarity 43.1%; Pred. No. 4.4e-42;  
 Matches 149; Conservative 71; Mismatches 107; Indels 19; Gaps 7;  
 QY 5 LVIIICIIAIGNCGLPLMLRYFNNGKRIWPFSTFLETAGPFPVIFIPLLFSYITRRSN 64  
 DB 49 LVIFSACLVLG-AGPPLLRLVYFVGGTFLMSATLQISGWPELLPLCVS-LYGRRH 106  
 QY 65 NVGDSTSFLLKRLI-AAVIGIISGFNDYLYAYGIAYLPVSTAAIIASQLAFIAIF 123  
 DB 107 GIGN----LLPRLVGAALVGLYAVSCFYALGSOALPLSTSLATQLATAFV 161  
 QY 124 SPMVKHKFPFFINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVAAAVMVF 183  
 DB 162 AFLFVGLRFFFSANAVLLTIGPVLGVGSPGKAGSSRAYWTGFCEAICAAALAGL 221  
 QY 184 ILPLVELAYOK-----AKQMTSYTLVLEFQLIILCLLASIVSVTGMFIAGDFKQALPK 235  
 DB 222 VIPLVEVATARYGRTGPAARPPPYATVMQAVMGAAGTAVCVLGMALKGDF-QAVAR 280  
 QY 236 EAREFKLGEALFYVAVFSAILWQFGLGALIFSTSSILVSGIMISVLLPITEVLAVIF 295  
 DB 281 EAAAFGLGAANYILVLAWDVSWOLLNGLMGLITCASSLLAGIMIAVLLPLSQVLAVIF 340  
 QY 296 YHEKFOAEKGLSLSLWGFVSFYGEIKSGEDKRRIQOESQETE 341  
 DB 341 LHEKPDGTGIALVLSLWGFASYLYGE--KAQKKEAQKMREREQE 384

RESULT 7

O49726 PRELIMINARY; PRT; 344 AA.  
 AC O49726; (TREMRELrel. 06, Created)  
 DT 01-JUN-1998 (TREMRELrel. 06, Last sequence update)  
 DT 01-OCT-2003 (TREMRELrel. 25, Last annotation update)  
 DE Hypothetical protein.  
 GN T9A21.70 OR AT4G18220.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]

SEQUENCE FROM N.A.  
 RP Bevan M., Murphy G., Drost L., Hall C., Hudson S., Ridley P.,  
 RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;  
 RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RA EU Arabidopsis sequencing project;  
 RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
 DR EMBL; AL021713; CAA16794.1; -;  
 DR EMBL; AL161548; CAB78824.1; -;  
 DR PIR; T04924; T04924.  
 DR InterPro; IPR004853; DUF250.  
 DR Pfam; PF03151; DUF250; 1.  
 KW Hypothetical protein.  
 SQ SEQUENCE 344 AA; 37819 MW; 28795F066738D8B1 CRC64;  
 Query Match 31.7%; Score 557.5; DB 10; Length 344;  
 Best Local Similarity 35.3%; Pred. No. 1.7e-32;  
 Matches 120; Conservative 72; Mismatches 131; Indels 17; Gaps 4;  
 QY 12 ILAIGNCGPLMLRYFNNGKRIWPFSTFLETAGPFPVIFIPLLFSYITRRSNVGDST 71  
 DB 7 VQVIGQSVATILGLRYENGSGSKMLATVQLVGFPIILLPYHLLSVKTHTTTQRDGLTS 66  
 QY 72 FFLIKPRLIIAAVIGIISGFNDYLYAYGIAYLPVSTAAIIASQLAFIAIFSFVVKHK 131  
 DB 67 ---LRNRALY-YIVLGLVGAACLYSIGLLYLPVSTLISLICASQLAFAFFSYLNSQK 122  
 QY 132 FTPTINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVAAAVMVFILPLVELA 191  
 DB 123 LTPILNLSFELLITSSLLAFNNEEDSKVTKGVKGVCTVGVASAGFGLLSQLA 182  
 QY 192 YQKAKQMTSYTLVLEFQLIILCLLASIVSVTGMFIAGDFKQALPKAREFKLGEALFYVVA 251  
 DB 183 FRKVLKQTFSEVINMIYMSLVASCVSVVGLFASSEWK-TLSEMEYIKLGKVSVMNL 241  
 QY 252 VFSAILWQFGLGALIFSTSSILVSGIMISVLLPITEVLAVIFYHEKFOAEKGLSLALS 311  
 DB 242 VMTAVTVQVFSIGCTGLIFELSSLSFSAISALGLPVVILAVIFHDKMGLKVISMILA 301  
 QY 312 LWGFVSFY-----GETKSGEDKRRIQOESQE 339  
 DB 302 IWGFVSYYVQYLDLTKKSNIEPTTSPDRPEAGSSE 341  
 RESULT 8  
 Q8RY74 PRELIMINARY; PRT; 377 AA.  
 AC Q8RY74; (TREMRELrel. 21, Created)  
 DT 01-JUN-2002 (TREMRELrel. 21, Last sequence update)  
 DT 01-OCT-2002 (TREMRELrel. 22, Last annotation update)  
 DE Hypothetical protein.  
 GN AT4G18200.  
 OS Arabidopsis thaliana (Mouse-ear cress).  
 OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
 OC eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
 OX NCBI\_TaxID=3702;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,  
 RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,  
 RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,  
 RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,  
 RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Satou M.,  
 RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,  
 RA Theologis A.;  
 RT "Arabidopsis Full Length cDNA Clones.";





```

RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.M., Hayashizaki Y., Heuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Narusaka M., Onodera C.S., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Shinn P., Tang C.C., Toroumi M., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinzaki K., Ecker J., Theologis A.,
RA Davis R.W.
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL: AC020576; AAF78258.1; -.
DR EMBL: AY062782; AAL32860.1; -.
DR EMBL: BT003358; AAO29976.1; -.
DR FIR: D96506; D96506.
DR InterPro: IPR004853; DUF250.
DR Pfam: PF03151; DUF250; 1.
KW Hypothetical protein.
SQ SEQUENCE 379 AA; 41289 MW; F1ED85A70FAD7001 CRC64;
Query Match 31.6%; Score 556.5; DB 10; Length 379;
Best Local Similarity 34.4%; Pred. No. 2.3e-32;
Matches 115; Conservative 89; Mismatches 121; Indels 9; Gaps 4;
QY 5 LVINCIILAIAGCGGLIMRLFYFNGKRIWFSTFLETAGFPVPIPLFSYITRRSN 64
DB 45 LVSVINFFLIGQASVLLGRFYDEGGNSKMATLVQTAAPPIIYIPLL----LPSSA 100
QY 65 NVGDSFPLIKRLLIAVTVIGLSGPDNYLYANGIYAVLPVSTAAIIASQAFIAIFS 124
DB 101 SVESSESSCLYIVLI-VYLGVIITAGDNMLYSVGLLYLSASTVSLICATQAFNAVFS 159
QY 125 FFMVHKFPFTINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVAAAVYAFI 184
DB 160 YFINAQKFTALIINSVLLSFSALANDADTPSGVSRKYIVGVFCTLAASALYSLL 219
QY 185 LPILVELAYOKAKOTMSTYTLVLEFQLILLASIVSGNFIAGDPKQALPKAREFKLGE 244
DB 220 LSLMQSFKEILKRETFVSVMLEQIYTSVATCVSVIGLFSAGSEWR-TLHGEMEGYHKQ 278
QY 245 ALFVVAVFSAITWQGFGLGALGLIFSTSLVSGIMISVLLPITEVLAVIFVHEKFOAEK 304
DB 279 ASVYTLVWTAVTWQCVSGVGVGLIIFLVTSLSFNVISTISLAVTPLAALVVFDDKMGVK 338
QY 305 GLSLALSLWGFVSFYGEIKSGDKRIQEEQ 338
DB 339 IMAMLIAWGFASVYV---QNHIDDLKVRQARQQ 369
RESULT 11
ID O49725 PRELIMINARY; PRT; 348 AA.
AC O49725;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T9A21.60 OR AT4G18210.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurossids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Drost L., Hall C., Hudson S., Ridley P.,
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL: AL021713; CAA16793.1; -.
DR EMBL: AL161548; CAB78823.1; -.
DR FIR: T04923; T04923.
DR InterPro: IPR004853; DUF250.
DR Pfam: PF03151; DUF250; 1.
KW Hypothetical protein.
SQ SEQUENCE 348 AA; 38556 MW; 354CB8D6D5547786 CRC64;
Query Match 31.6%; Score 555.5; DB 10; Length 348;
Best Local Similarity 36.7%; Pred. No. 2.5e-32;
Matches 120; Conservative 72; Mismatches 120; Indels 15; Gaps 6;
QY 26 LYPNGGKRIWFSTFLETAGFPVPIPLFSYITRRSNVNGDSTSFLLIKPR-LLIAAV 84
DB 25 VYVDNGSGKWLATVQLVGFVLLPYFYLSPKTHATTDGRKTS-----PRNRVIVV 79
QY 85 IVGILSGFDNYLYANGIYAVLPVSTAAIIASQAFIAIFSFFMVHVKHFTPTINAVLLT 144
DB 80 VLGLLVGADCYLSIGLLYLPVSTYSLICASQAFNAVFSYFLNSQKLTPIILNSFLLT 139
QY 145 VGAALVGMHTETDKPVHETHKQYITGFLITVAAAVYAFILPLVELAYOKAKOTMSTYTLV 204
DB 140 ISSTLLAFNNEETDSTKVTGKGYVKGFICTVAASAGYGLVLSLQQLAFLLKQKQNFSEV 199
QY 205 LEFQLILLASIVSGNFIAGDPKQALPKAREFKLGEALFYVAVFSAIIWQGFLLG 264
DB 200 MDMIIVSVLASCVVVGLFASSEWK-TLSSEMDNVKHGKVSVMNLVMTAVTWQVFSIG 258
QY 265 AIGLIFSTSLVSGIMISVL-LPITEVLAVIFVHEKFOAEKGLSLALSFWGFSYFYGE- 322
DB 259 GTGLIFELSLFSN-AISVLGLPWPVILAVIIPHDKNGLKVLISMILAIWGFSTSYVYQ 317
QY 323 -----IKSGDKRIQEESEQTEQSS 344
DB 318 LDDKNLKNKHEITTTESPPDPEABEST 344
RESULT 12
QYX7P4 PRELIMINARY; PRT; 751 AA.
ID Q7X7P4
AC Q7X7P4;
DT 01-OCT-2003 (TrEMBLrel. 25, Created)
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE OSUNBA0013K16.25 protein [OSUNBB0016D16.5 protein].
GN OSUNBA0013K16.25 OR OSUNBB0016D16.5.
OS Oryza sativa (Rice).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartioideae; Oryzeae; Oryza.
OX NCBI_TaxID=4530;
RN [1]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.

```



RP	SEQUENCE FROM N.A.	
RA	Cheuk R., Shinn P., Brooks S., Buehler E., Chao Q., Johnson-Hopson C.,	
RA	Khan S., Kim C., Altafi H., Bei B., Chin C., Chiou J., Choi E.,	
RA	Conn L., Conway A., Gonzalez A., Hansen N., Howing B., Koo T., Lam B.,	
RA	Lee J., Lenz C., Li J., Liu A., Liu J., Liu S., Mukharsky N.,	
RA	Nguyen M., Palm C., Pham P., Sakano H., Schwartz J., Southwick A.,	
RA	Thaveri A., Toriumi M., Vaysberg M., Yu G., Davis R., Federspiel N.,	
RA	Theologis A., Ecker J.;	
RL	Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.	
RN	[6]	
RP	SEQUENCE FROM N.A.	
RA	Haas B.J., Volforsky N., Town C.D., Troukhan M., Alexandrov N.,	
RA	Feldmann K.A., Flavell R.B., White O., Salzberg S.B.;	
RT	"Full-length messenger RNA sequences greatly improve genome	
RT	annotation.";	
RT	Genome Biol. 0:0-0 (2002).	
RN	[7]	
RP	SEQUENCE FROM N.A.	
RA	Brover V., Troukhan M., Alexandrov N., Lu Y.-P., Flavell R.,	
RA	Feldmann K.;	
RT	"Full-length cDNA from Arabidopsis thaliana.";	
RL	Submitted (MAR-2002) to the EMBL/GenBank/DBJ databases.	
RN	[8]	
RP	SEQUENCE FROM N.A.	
RA	Yamada K., Chan M.M., Chang C.H., Dale J.M., Huan V.W., Lee J.M.,	
RA	Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,	
RA	Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,	
RA	Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,	
RA	Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,	
RA	Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,	
RA	Theologis A.;	
RT	"Arabidopsis Full Length cDNA Clones.";	
RL	Submitted (FEB-2003) to the EMBL/GenBank/DBJ databases.	
RN	[9]	
RP	SEQUENCE FROM N.A.	
RA	Yamada K., Chan M.M., Chang C.H., Dale J.M., Huan V.W., Lee J.M.,	
RA	Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,	
RA	Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,	
RA	Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,	
RA	Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,	
RA	Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,	
RA	Theologis A.;	
RT	"Arabidopsis Open Reading Frame (ORF) Clones.";	
RL	Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AC004135; AAD32940.1; -	
DR	EMBL; AY084641; AAM61204.1; -	
DR	EMBL; BT004160; AAO42180.1; -	
DR	EMBL; BT005495; AAO63915.1; -	
DR	Go; GO:0016020; C:membrane; IEA.	
DR	InterPro; IPR004853; DUF250.	
DR	InterPro; IPR000620; DUF6.	
DR	Pfam; PF03151; DUF250; 1.	
DR	Pfam; PF00892; DUF6; 1.	
KW	Hypothetical protein	
SQ	SEQUENCE 382 AA; 42320 MW; E030A23A8C834ACA CRC64;	
	Query Match 29.3%; Score 516.5; DB 10; Length 382;	
	Best Local Similarity 32.2%; Pred. No. 1.7e-29;	
	Matches 115; Conservative 81; Mismatches 132; Indels 29; Gaps 7;	
QY	1 MWK-----ALVINCIILAIIGNCGPLIMRLYFNNGKRIWFSTFLETAGFPVIFIP 53	
DB	16 MVKPPVKRSITLIVTFYFFGSIASSLLAKYFYVGGSSRWVSTWQSGAFPLLIIL 75	
QY	54 LFSY----ITRRSNNVGDSSTFELIKPRLLIAAVIVGILSGFDNYLYAYGIAYLPVSTA 109	
DB	76 YPHYVLKITTTR-----PPTRTL---RHLIFSLVIGLVGFNPNFSLWGTSYLPVSTS 127	
QY	110 ALIIAQLAFIAIFSPFMVKHKTPTPTINAVLLTVGAAVLGMHETDKPVHETHKQYIT 169	
DB	128 SLLLSQLVFTLILSRIVKQKITFNSLNCNVLTLSSVLLADSDSKDPESGLTKTKYFI 187	
QY	170 GFLITVAAAVYAFILPLVELAYOKAKQMTSYTLVLEFQLILCLLASIVSVIGMFTAGDF 229	
DB	188 GYSTTICAGLLFALYLPVTEKLY--RTVVCYAMVMEVQLVMEFAATVFATIGMACEGGF 244	
QY	230 KQALPKAEKREFKGEALFYVVAVFSAILI-WQFFLGAIGLIFSTSLVSGIMISVLLPIT 288	
DB	245 KEMVKEANHVFTKGPFTFYMTFAILANVVTQLSFAATSGMVYLTSGITGICMTALLAMN 304	
QY	289 EVLAVIFYHEKFOAEKGLSLALSLWGFVSFYFGEIKSGEDKRIIOQESQETPQSSL 345	
DB	305 VIGGVAYGDFGVGKIVSTVLCINGFSSYTYGMV-----MKMKKEEKGEYSVG 355	
	RESULT 15	
O49722	PRELIMINARY; PRT; 358 AA.	
ID	O49722;	
AC	O49722;	
DT	01-JUN-1998 (TrEMBLrel. 06, Created)	
DT	01-JUN-1998 (TrEMBLrel. 06, Last sequence update)	
DT	01-OCT-2003 (TrEMBLrel. 25, Last annotation update)	
DE	Hypothetical protein.	
GN	T9AZ1.30 OR A14G18190	
OS	Arabidopsis thaliana (Mouse-ear cress).	
OC	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;	
OC	Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;	
OC	eurosid II; Brassicales; Brassicaceae; Arabidopsi.	
OX	NCBI_TaxID=3702;	
RN	[1]	
RP	SEQUENCE FROM N.A.	
RA	Bevan M., Murphy G., Drost L., Hall C., Hudson S., Ridley P.,	
RA	Bancroft I., Mewes H.W., Mayer K., Schueller C.;	
RA	Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.	
RN	[2]	
RP	SEQUENCE FROM N.A.	
RA	Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;	
RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
RN	[3]	
RP	SEQUENCE FROM N.A.	
RA	EU Arabidopsis sequencing project;	
RA	Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.	
DR	EMBL; AL021713; CAAL6790.1; -	
DR	EMBL; AL161548; CAB78821.1; -	
DR	PIR; T04921; T04921.	
DR	InterPro; IPR004853; DUF250.	
DR	Pfam; PF03151; DUF250; 1.	
KW	Hypothetical protein.	
SQ	SEQUENCE 358 AA; 39552 MW; FDD45970C1CF1600 CRC64;	
	Query Match 28.6%; Score 502.5; DB 10; Length 358;	
	Best Local Similarity 33.2%; Pred. No. 1.7e-28;	
	Matches 113; Conservative 80; Mismatches 120; Indels 27; Gaps 7;	
QY	12 ILAIGNCGPLIMRLYFNNGKRIWFSTFLETAGFPVIFIPFLLPSYITRRSNNVGDS 71	
DB	29 LLLAGETIATLLGLRYEKGKSTWLETVLQLE-----PSKTKTITKK-----TTS 75	
QY	72 FFLIKPRLLIAAVIVGILSGFDNYLYAYGIAYLPVSTAALIIASQLAFIAIFSPFMVKH 131	
DB	76 SFL---TLSIVYIGLLVAGHCILYSGLLYLPVSTFSLISASQLAENAVFSYFLNSQK 132	
QY	132 PTPFTINAVLLTVGAAVLGMHETDKPVHETHK-----QYITGFLITVAAAVYAFILP 186	
DB	133 ITPFTINSLVLLTISSTLLVIQHEPESP-SSTSKSAKSKYVIGYICAVGSAGSYSLVS 191	
QY	187 LVELAYOKAKQMTSYTLVLEFQLILCLLASIVSVIGMFIAGDFKQALPKAEKREFKLGAL 246	
DB	192 LTDVAFEKILKKYTFKALIDMATVPWMATCVVVVGLFGSGWK-KLSTEMEELQKSS 250	
QY	247 FVVAVFSAILIQQGFPLGAILGFSTSLVSGIMISVLLPITEVLAVIFYHEKFOAEKGL 306	
DB	251 YILINIGTISQACLIGSVGLIIEVSLFSNVISTLCLPVVPVLAIVVFFRDEMGIKLV 310	
QY	307 SLALSLSWGFVSFYF---GEIKSGEDKRIIOQESQETQ 342	

Db 311 AMFLAIWGFVSYGQHYVNDKPEEDQELFQSKKEEQKQ 350

Search completed: September 23, 2004, 23:49:47  
Job time : 75.5819 secs

**THIS PAGE BLANK (USPTO)**





66	20	1.6	104654	2	AC141775	139	19	1.6	659	8	AF370226	AF370226 Arabidops
67	20	1.6	108365	8	AC015447	140	19	1.6	717	8	WHTPH2AA	D38087 Triticum ae
68	20	1.6	112066	9	AC007027	141	19	1.6	738	6	AX865261	AX865261 Sequence
69	20	1.6	120423	9	AC084752	142	19	1.6	738	6	BD145323	BD145323 Primer fo
70	20	1.6	132229	9	HS447N6	c 143	19	1.6	938	3	AF117590	AF117590 Manduca s
71	20	1.6	134028	9	AL662906	144	19	1.6	1029	11	CNS061I13	AL400321 T7 end of
72	20	1.6	151616	2	AC121662	145	19	1.6	1067	6	AX033553	AX033553 Sequence
73	20	1.6	152302	2	AC123898	146	19	1.6	1319	6	AX509845	AX509845 Sequence
74	20	1.6	154336	9	AP006300	147	19	1.6	1319	6	AX652563	AX652563 Sequence
75	20	1.6	155262	10	AL592489	148	19	1.6	1696	6	AX427518	AX427518 Sequence
76	20	1.6	155735	2	AC145288	149	19	1.6	1701	8	EMECMKA	M74120 Emericella
77	20	1.6	156524	9	AC136305	c 150	19	1.6	1748	8	AF313491	AF313491 Matthiola
78	20	1.6	157653	9	AC005146	151	19	1.6	1753	8	AK067521	AK067521 Oryza sat
79	20	1.6	157904	9	AC108486	152	19	1.6	1854	9	BC005368	BC005368 Homo sapi
80	20	1.6	160373	10	AC128291	153	19	1.6	1857	8	AK118410	AK118410 Arabidops
81	20	1.6	160610	2	AC115899	154	19	1.6	1920	10	RNFBASE2	X83735 R. norvegicu
82	20	1.6	160626	10	AL833796	c 155	19	1.6	2060	6	AX875567	AX875567 Sequence
83	20	1.6	161177	2	AC132338	c 156	19	1.6	2060	6	BD155628	BD155628 Primer fo
84	20	1.6	161677	2	AC125850	c 157	19	1.6	2060	6	AX000865	AX000865 Homo sapi
85	20	1.6	162325	9	AC109585	158	19	1.6	2082	8	AY145024	AY145024 Saccharom
86	20	1.6	163209	9	AC093419	159	19	1.6	2132	6	AX875653	AX875653 Sequence
87	20	1.6	163509	9	AC017015	160	19	1.6	2132	6	BD155683	BD155683 Primer fo
88	20	1.6	164689	2	AL953873	161	19	1.6	2132	9	AK000909	AK000909 Homo sapi
89	20	1.6	165126	9	CNS01RHV	162	19	1.6	2730	9	HUMSNF2LB	M89907 Human globa
90	20	1.6	166462	2	AC027064	163	19	1.6	2879	8	AF054580	AF054580 Emericell
91	20	1.6	169548	2	AC119703	164	19	1.6	2898	9	BC051825	BC051825 Homo sapi
92	20	1.6	170269	9	AC012085	165	19	1.6	3097	9	AK026426	AK026426 Homo sapi
93	20	1.6	170674	2	AC141744	166	19	1.6	3176	9	HSMB03019	AL713677 Homo sapi
94	20	1.6	174285	2	AC145484	167	19	1.6	3424	9	BC034228	BC034228 Homo sapi
95	20	1.6	175326	2	AL954772	168	19	1.6	3634	6	AX409769	AX409769 Sequence
96	20	1.6	177552	2	BX530015	169	19	1.6	3634	9	AX822007	AX822007 Sequence
97	20	1.6	177788	2	AC069310	170	19	1.6	3634	6	HUMSNF2L	M88163 Human globa
98	20	1.6	177911	2	AC021362	c 171	19	1.6	3647	6	AR068838	AR068838 Sequence
99	20	1.6	181541	9	AL136524	c 172	19	1.6	3647	6	AR070138	AR070138 Sequence
100	20	1.6	182912	9	AC068060	c 173	19	1.6	3653	10	RATHEXOKIN	J04526 Rattus norv
101	20	1.6	182913	9	AC068556	c 174	19	1.6	3717	8	AB026726	AB026726 Pichia an
102	20	1.6	183179	2	AC101922	c 175	19	1.6	4376	9	HSMB02695	AL355937 Homo sapi
103	20	1.6	184330	9	AC093561	176	19	1.6	6419	6	AX323764	AX323764 Sequence
104	20	1.6	196537	2	AC073293	177	19	1.6	6419	6	AX346261	AX346261 Sequence
105	20	1.6	198614	2	AC073293	c 178	19	1.6	9211	3	AB030033	AB030033 Dictyoste
106	20	1.6	200098	5	AL935317	179	19	1.6	10205	6	AX344270	AX344270 Sequence
107	20	1.6	201728	10	AC011013	180	19	1.6	10205	6	AX348667	AX348667 Sequence
108	20	1.6	202998	2	AC072058	c 181	19	1.6	12090	9	AB057594	AB057594 Homo sapi
109	20	1.6	213050	1	AL646067	c 182	19	1.6	20162	10	AL772322	AL772322 Mouse DNA
110	20	1.6	217293	10	AC099584	c 183	19	1.6	24267	3	CEK08D8	W73104 Caenorhabdi
111	20	1.6	219796	2	AC146734	c 184	19	1.6	37656	3	AC006608	AC006608 Caenorhab
112	20	1.6	224005	2	AC120997	c 185	19	1.6	39667	10	AF1731863	AF1731863 Mouse DNA
113	20	1.6	227433	2	AC126076	c 186	19	1.6	40143	10	AF185591	AF185591 Mus muscu
114	20	1.6	227774	2	AC120894	c 187	19	1.6	40951	3	AF067211	AF067211 Caenorhab
115	20	1.6	232571	10	AC116557	c 188	19	1.6	40951	3	AF067211	AF067211 Caenorhab
116	20	1.6	233550	2	AC140761	c 189	19	1.6	42722	2	AC100936	AC100936 Homo sapi
117	20	1.6	235115	2	AC133256	c 190	19	1.6	52454	2	AC020246	AC100936 Mus muscu
118	20	1.6	236561	2	BX897697	c 191	19	1.6	5454	2	AC020246	AC020246 Drosophil
119	20	1.6	239880	2	AC106296	c 192	19	1.6	56822	2	AC099976	AC099976 Bos tauru
120	20	1.6	242829	2	AC114391	c 193	19	1.6	6102	2	AC129953	AC129953 Bos tauru
121	20	1.6	244047	10	AC107780	c 194	19	1.6	67224	2	AC129953	AC129953 Bos tauru
122	20	1.6	249068	2	AC095295	c 195	19	1.6	67224	2	AC109287	Continuation (4 of
123	20	1.6	250029	9	AE014307	c 196	19	1.6	68487	2	AL138745	AL138745 Human DNA
124	20	1.6	255044	2	AC118392	c 197	19	1.6	73632	10	AC098685	AC098685 Mus muscu
125	20	1.6	261330	2	AC103472	c 198	19	1.6	76798	9	HS32281	Z84718 Human DNA s
126	20	1.6	263129	2	AC103472	c 199	19	1.6	76798	9	MM41M10	AL136328 Mus muscu
127	20	1.6	263911	2	AC095364	c 200	19	1.6	78785	10	MM41M10	AP004159 Oryza sat
128	20	1.6	263954	2	AC117122	c 201	19	1.6	80820	8	AP004159	AP004159 Oryza sat
129	20	1.6	266390	2	AC106133	c 202	19	1.6	81583	8	AP004159	AC131927 Homo sapi
130	20	1.6	269081	2	AC068493	c 203	19	1.6	81729	8	AB016874	AB016874 Arabidops
131	20	1.6	326890	2	AC135913	c 204	19	1.6	82359	8	ATAC002342	AC002342 Arabidops
132	19	1.6	243	14	HIVGIM405	c 205	19	1.6	82584	8	NCB13B3	NCB13B3 Neurospor
133	19	1.6	435	6	AX765870	c 206	19	1.6	82966	5	AC140932	AC140932 Gallus ga
134	19	1.6	466	8	AY059118	c 207	19	1.6	86719	9	AP000885	AP000885 Mus muscu
135	19	1.6	500	6	AR423859	c 208	19	1.6	89782	2	AC145298	AC145298 Homo sapi
136	19	1.6	500	6	BD119412	c 209	19	1.6	90476	9	AL015310	AL015310 Homo sapi
137	19	1.6	589	11	BV077572	c 210	19	1.6	90631	10	AL845350	AL845350 Mouse DNA
138	19	1.6	644	8	AY087711	c 211	19	1.6	92612	8	AC003974	AC003974 Arabidops

212	19	1.6	93964	2	AC022665	AC022665 Homo sapi	19	1.6	156228	10	AC128768	AC128768 Rattus no
c 213	19	1.6	93964	2	AC022665	AC022665 Homo sapi	19	1.6	156263	8	AC137614	Oryza sat
214	19	1.6	95888	2	AC136845	AC136845 Rattus no	19	1.6	156357	2	CNS0984X	AC137614 Oryza sat
215	19	1.6	99460	2	AC106346	Continuation (4 of	19	1.6	156357	2	CNS0984X	AC137614 Oryza sat
c 216	19	1.6	100343	10	BS571735	Mouse DNA	19	1.6	156507	2	AC117859	AC117859 Rattus no
c 217	19	1.6	100685	8	AC023754	Arabidops	19	1.6	156507	2	AC117859	AC117859 Rattus no
c 218	19	1.6	101295	3	CEY488B6A	Alu10490 Caenorhab	19	1.6	156628	2	AC146761	AC146761 Pan trogl
c 219	19	1.6	101824	3	AL518253	Human DNA	19	1.6	157058	9	CNS05TBY	AC157058 Homo sapi
c 220	19	1.6	102880	9	AL512403	Human DNA	19	1.6	157470	2	CNS05TBY	AC157058 Homo sapi
c 221	19	1.6	103353	8	AL512403	Human DNA	19	1.6	157470	2	CNS05TBY	AC157058 Homo sapi
c 222	19	1.6	104607	8	AT728J14	Arabidops	19	1.6	157834	2	AC022931	AC022931 Homo sapi
c 223	19	1.6	105223	8	AT728J14	Arabidops	19	1.6	157834	2	AC022931	AC022931 Homo sapi
c 224	19	1.6	105622	2	AC137517	Homo sapi	19	1.6	158094	9	AC034245	AC034245 Homo sapi
c 225	19	1.6	107381	2	AC137517	Homo sapi	19	1.6	158094	9	AC034245	AC034245 Homo sapi
c 226	19	1.6	107705	3	AC024826	Caenorhab	19	1.6	158420	9	AC055855	AC055855 Homo sapi
c 227	19	1.6	108950	2	AC135231	Medicago	19	1.6	159287	9	AP003694	AP003694 Homo sapi
c 228	19	1.6	109929	10	AC093922	Genomic s	19	1.6	159833	2	AC105376	AC105376 Pelis cat
c 229	19	1.6	110000	2	AC115434	Rattus no	19	1.6	160779	10	AL732449	AL732449 Mouse DNA
c 230	19	1.6	110000	2	AC115434	Rattus no	19	1.6	160779	10	AL732449	AL732449 Mouse DNA
c 231	19	1.6	110000	2	AC115434	Rattus no	19	1.6	162065	2	AC117061	AC117061 Rattus no
c 232	19	1.6	110000	2	AC115434	Rattus no	19	1.6	162065	2	AC117061	AC117061 Rattus no
c 233	19	1.6	111698	2	AC113066	Mus muscu	19	1.6	162247	2	AC019845	AC019845 Drosophil
c 234	19	1.6	112184	9	AC008805	Homo sapi	19	1.6	163378	10	AC133182	AC133182 Mus muscu
c 235	19	1.6	112235	1	AE012561	Xylella f	19	1.6	163378	10	AC133182	AC133182 Mus muscu
c 236	19	1.6	112785	10	AL929332	Mouse DNA	19	1.6	163432	2	BS511091	BS511091 Danio rer
c 237	19	1.6	113160	2	AC144893	Medicago	19	1.6	163432	2	BS511091	BS511091 Danio rer
c 238	19	1.6	113754	9	HSJ342K12	Human DNA	19	1.6	163932	2	AC146625	AC146625 Papio anu
c 239	19	1.6	118049	8	AC104276	Oryza sat	19	1.6	163932	2	AC146625	AC146625 Papio anu
c 240	19	1.6	118196	8	ATF18A5	Arabidops	19	1.6	164526	9	AC114969	AC114969 Homo sapi
c 241	19	1.6	118999	9	AF240786	Homo sapi	19	1.6	164526	9	AC114969	AC114969 Homo sapi
c 242	19	1.6	118999	9	AF240786	Homo sapi	19	1.6	164864	2	AC115059	AC115059 Mus muscu
c 243	19	1.6	119065	2	AC136141	Medicago	19	1.6	164864	2	AC115059	AC115059 Mus muscu
c 244	19	1.6	120007	9	AF064864	Homo sapi	19	1.6	164864	2	AC115059	AC115059 Mus muscu
c 245	19	1.6	120557	9	AF064864	Homo sapi	19	1.6	164915	2	AC141422	AC141422 Pan trogl
c 246	19	1.6	120609	9	AF165176	Homo sapi	19	1.6	165237	2	AC138665	AC138665 Mus muscu
c 247	19	1.6	121556	8	CNS08CD2	Medicago	19	1.6	165237	2	AC138665	AC138665 Mus muscu
c 248	19	1.6	121721	9	AC090195	Homo sapi	19	1.6	165913	2	AC091914	AC091914 Homo sapi
c 249	19	1.6	127286	2	AC025286	Homo sapi	19	1.6	165913	2	AC091914	AC091914 Homo sapi
c 250	19	1.6	129413	9	HSJ709L21	Human DNA	19	1.6	166337	10	AL805982	AL805982 Mouse DNA
c 251	19	1.6	129790	2	CNS08CBJ	Oryza sat	19	1.6	166337	10	AL805982	AL805982 Mouse DNA
c 252	19	1.6	129790	2	CNS08CBJ	Oryza sat	19	1.6	166937	2	AC116927	AC116927 Rattus no
c 253	19	1.6	130743	2	AC122726	Medicago	19	1.6	166937	2	AC116927	AC116927 Rattus no
c 254	19	1.6	131697	10	AL731694	Mouse DNA	19	1.6	166992	2	AC119973	AC119973 Mus muscu
c 255	19	1.6	133289	9	AC012317	Homo sapi	19	1.6	166992	2	AC119973	AC119973 Mus muscu
c 256	19	1.6	133624	10	AC125226	Mus muscu	19	1.6	167541	2	AC102288	AC102288 Mus muscu
c 257	19	1.6	135578	2	AC117640	Mus muscu	19	1.6	167541	2	AC102288	AC102288 Mus muscu
c 258	19	1.6	137493	10	AL732581	Mouse DNA	19	1.6	167699	9	AC087379	AC087379 Homo sapi
c 259	19	1.6	142152	2	AC139658	Rattus no	19	1.6	167699	9	AC087379	AC087379 Homo sapi
c 260	19	1.6	143969	9	AL357552	Human DNA	19	1.6	167768	9	AC026224	AC026224 Homo sapi
c 261	19	1.6	144988	10	AC101022	Mus muscu	19	1.6	167768	9	AC026224	AC026224 Homo sapi
c 262	19	1.6	145035	9	CNS07EEY	Human DNA	19	1.6	168080	9	AC073172	AC073172 Homo sapi
c 263	19	1.6	145668	9	AC092272	Homo sapi	19	1.6	168080	9	AC073172	AC073172 Homo sapi
c 264	19	1.6	146113	2	AC125851	Rattus no	19	1.6	168603	2	BS571842	BS571842 Danio rer
c 265	19	1.6	146715	9	AL355795	Human DNA	19	1.6	168603	2	BS571842	BS571842 Danio rer
c 266	19	1.6	146832	10	AL645844	Mouse DNA	19	1.6	168730	10	AC121597	AC121597 Mus muscu
c 267	19	1.6	147280	9	AC067747	Homo sapi	19	1.6	168730	10	AC121597	AC121597 Mus muscu
c 268	19	1.6	147804	10	AL645647	Mouse DNA	19	1.6	169106	2	AC025240	AC025240 Homo sapi
c 269	19	1.6	147941	8	AC123525	Oryza sat	19	1.6	169106	2	AC025240	AC025240 Homo sapi
c 270	19	1.6	148101	2	BS530071	Danio rer	19	1.6	170048	2	AC025240	AC025240 Homo sapi
c 271	19	1.6	149267	9	AC087699	Homo sapi	19	1.6	170048	2	AC025240	AC025240 Homo sapi
c 272	19	1.6	149602	2	AC137291	Rattus no	19	1.6	170356	3	AC010558	AC010558 Drosophil
c 273	19	1.6	149879	10	AL591429	Mouse DNA	19	1.6	170356	3	AC010558	AC010558 Drosophil
c 274	19	1.6	149980	2	AC116836	Mus muscu	19	1.6	170980	2	AF277371	AF277371 Homo sapi
c 275	19	1.6	151447	3	AC010122	Drosophil	19	1.6	171682	9	AL353072	AL353072 Human DNA
c 276	19	1.6	152869	2	AC132275	Rattus no	19	1.6	171682	9	AL353072	AL353072 Human DNA
c 277	19	1.6	153763	9	AC016933	Homo sapi	19	1.6	172071	9	AC012676	AC012676 Homo sapi
c 278	19	1.6	155185	9	AC023549	Homo sapi	19	1.6	172071	9	AC012676	AC012676 Homo sapi
c 279	19	1.6	155263	5	BS005418	Zebrafish	19	1.6	172123	10	AC060809	AC060809 Homo sapi
c 280	19	1.6	155448	2	AC024053	Homo sapi	19	1.6	172123	10	AC060809	AC060809 Homo sapi
c 281	19	1.6	155760	2	AC010076	Homo sapi	19	1.6	172714	2	AC112269	AC112269 Mus muscu
c 282	19	1.6	155813	2	AC019350	Homo sapi	19	1.6	172714	2	AC112269	AC112269 Mus muscu
c 283	19	1.6	155881	2	AC102713	Mus muscu	19	1.6	173316	2	AC141726	AC141726 Apis mell
c 284	19	1.6	156077	9	AL354997	Human DNA	19	1.6	173316	2	AC141726	AC141726 Apis mell

358	19	1.6 183132	2	AC119637	AC119637 Rattus no	431	19	1.6 213330	2	AC133321	AC133321 Rattus no
359	19	1.6 183912	2	AC021240	AC021240 Homo sapi	C 432	19	1.6 213362	2	AC142467	AC142467 Rattus no
360	19	1.6 184174	10	AC121570	AC121570 Mus muscu	C 433	19	1.6 213870	10	AL844180	AL844180 Mouse DNA
361	19	1.6 184352	9	BS000179	BS000179 Pan trogl	C 434	19	1.6 213959	2	AC102642	AC102642 Mus muscu
362	19	1.6 184767	2	AC001948	AC001948 Mus muscu	435	19	1.6 214306	2	AC128723	AC128723 Rattus no
363	19	1.6 185109	2	AC141546	AC141546 Rattus no	436	19	1.6 214340	2	AC130012	AC130012 Rattus no
364	19	1.6 185349	2	AC105173	AC105173 Mus muscu	C 437	19	1.6 214687	9	AC097636	AC097636 Homo sapi
365	19	1.6 185441	10	AL824716	AL824716 Mouse DNA	438	19	1.6 214793	2	AC023345	AC023345 Homo sapi
366	19	1.6 187214	2	AC137676	AC137676 Mus muscu	439	19	1.6 214993	2	AC055773	AC055773 Mus muscu
367	19	1.6 187582	2	AC139051	AC139051 Mus muscu	440	19	1.6 215106	10	AL607070	AL607070 Mouse DNA
368	19	1.6 187732	9	CNS01RHM	AL161851 Human chr	441	19	1.6 215178	2	AC132001	AC132001 Mus muscu
369	19	1.6 187779	9	AL359703	AL359703 Human DNA	C 442	19	1.6 215378	2	AC115851	AC115851 Mus muscu
370	19	1.6 188406	10	AL713914	AL713914 Mouse DNA	C 443	19	1.6 215516	2	AC111544	AC111544 Rattus no
371	19	1.6 188545	2	AC133900	AC133900 Mus muscu	C 444	19	1.6 215520	2	AC115764	AC115764 Mus muscu
372	19	1.6 188803	2	AC115454	AC115454 Rattus no	C 445	19	1.6 215563	2	AC111429	AC111429 Rattus no
373	19	1.6 188928	10	AC121814	AC121814 Mus muscu	C 446	19	1.6 216044	10	AC114824	AC114824 Mus muscu
374	19	1.6 189180	10	AC127235	AC127235 Mus muscu	447	19	1.6 216498	2	AC123722	AC123722 Mus muscu
375	19	1.6 189541	2	AC118520	AC118520 Rattus no	C 448	19	1.6 216746	2	AC133946	AC133946 Mus muscu
376	19	1.6 189541	2	AC144809	AC144809 Mus muscu	C 449	19	1.6 216937	2	AC022061	AC022061 Mus muscu
377	19	1.6 189712	9	AC023034	AC023034 Homo sapi	C 450	19	1.6 217097	2	AC112080	AC112080 Rattus no
378	19	1.6 189771	2	AC132188	AC132188 Danio rer	451	19	1.6 217248	2	AC124626	AC124626 Mus muscu
379	19	1.6 189909	2	AC131912	AC131912 Mus muscu	C 452	19	1.6 217522	10	AC117639	AC117639 Mus muscu
380	19	1.6 190349	9	AC009475	AC009475 Homo sapi	453	19	1.6 217558	2	AC125449	AC125449 Mus muscu
381	19	1.6 190627	2	AC132589	AC132589 Mus muscu	454	19	1.6 217790	2	AC079469	AC079469 Homo sapi
382	19	1.6 190712	10	AC100382	AC100382 Mus muscu	455	19	1.6 217990	2	AC115978	AC115978 Mus muscu
383	19	1.6 190724	2	AC111312	AC111312 Danio rer	456	19	1.6 218074	2	AC103155	AC103155 Rattus no
384	19	1.6 191392	2	AC133238	AC133238 Rattus no	457	19	1.6 218552	10	AC118607	AC118607 Mus muscu
385	19	1.6 192370	2	AC016857	AC016857 Homo sapi	C 458	19	1.6 218613	2	AC113601	AC113601 Mus muscu
386	19	1.6 192688	10	AL928805	AL928805 Mouse DNA	459	19	1.6 218665	2	AC109875	AC109875 Rattus no
387	19	1.6 193355	9	CNS01RH4	AL60471 Human chr	460	19	1.6 218865	2	AC096316	AC096316 Rattus no
388	19	1.6 194077	10	AC117221	AL672081 Mouse DNA	C 461	19	1.6 219412	10	AL772341	AL772341 Mouse DNA
389	19	1.6 194141	10	AC110737	AC110737 Mus muscu	C 462	19	1.6 219905	2	AC128936	AC128936 Rattus no
390	19	1.6 194168	2	AC111038	AC111038 Mus muscu	C 463	19	1.6 221047	10	AC123068	AC123068 Mus muscu
391	19	1.6 194546	2	AC139927	AC139927 Rattus no	C 464	19	1.6 221464	2	AC137905	AC137905 Mus muscu
392	19	1.6 195413	2	AC125895	AC125895 Rattus no	C 465	19	1.6 221539	2	AC098141	AC098141 Rattus no
393	19	1.6 195634	2	AL663045	AL663045 Mouse DNA	C 466	19	1.6 222845	10	AC010049	AC010049 Mus muscu
394	19	1.6 195656	10	AC109504	AC109504 Mus muscu	C 467	19	1.6 223797	2	AC021978	AC021978 Homo sapi
395	19	1.6 196649	9	AC007684	AC007684 Homo sapi	C 468	19	1.6 224003	2	AC021978	AC021978 Homo sapi
396	19	1.6 196848	9	AC137902	AC137902 Mus muscu	C 469	19	1.6 224649	2	AC109747	AC109747 Rattus no
397	19	1.6 196890	2	AC137902	AC137902 Mus muscu	C 470	19	1.6 224734	2	AC119172	AC119172 Rattus no
398	19	1.6 196976	2	AC137902	AC137902 Mus muscu	C 471	19	1.6 224979	2	AC114178	AC114178 Rattus no
399	19	1.6 197402	2	AC137902	AC137902 Mus muscu	C 472	19	1.6 225235	2	AC117100	AC117100 Rattus no
400	19	1.6 198051	2	AC137902	AC137902 Mus muscu	C 473	19	1.6 225258	5	AC114178	AC114178 Rattus no
401	19	1.6 198323	2	AC137902	AC137902 Mus muscu	C 474	19	1.6 225586	5	AC133121	AC133121 Rattus no
402	19	1.6 198323	2	AC137902	AC137902 Mus muscu	C 475	19	1.6 225626	2	AC133121	AC133121 Rattus no
403	19	1.6 199209	2	AC137902	AC137902 Mus muscu	C 476	19	1.6 225958	2	AC140794	AC140794 Rattus no
404	19	1.6 199224	2	AC137902	AC137902 Mus muscu	C 477	19	1.6 225978	2	AC140462	AC140462 Mus muscu
405	19	1.6 199667	8	ATC9HIV37	AC112347 Rattus no	C 478	19	1.6 225978	2	AC140462	AC140462 Mus muscu
406	19	1.6 200758	2	AC112347	AC112347 Rattus no	C 479	19	1.6 226759	2	AC114503	AC114503 Rattus no
407	19	1.6 201371	2	AC119162	AC119162 Mus muscu	C 480	19	1.6 227632	9	AC087382	AC087382 Homo sapi
408	19	1.6 201935	10	AC121582	AC121582 Mus muscu	C 481	19	1.6 227920	2	AC107510	AC107510 Rattus no
409	19	1.6 202689	10	AL672232	AL672232 Mouse DNA	C 482	19	1.6 228511	2	AC127653	AC127653 Rattus no
410	19	1.6 203641	2	AC139568	AC139568 Mus muscu	C 483	19	1.6 230615	2	AC095652	AC095652 Rattus no
411	19	1.6 203641	10	AL845297	AL845297 Mouse DNA	C 484	19	1.6 230615	2	AC131547	AC131547 Rattus no
412	19	1.6 203839	2	AC091288	AC091288 Mus muscu	C 485	19	1.6 232885	2	AC106163	AC106163 Rattus no
413	19	1.6 204136	10	AL772271	AL772271 Mouse DNA	C 486	19	1.6 232950	10	AL670236	AL670236 Mouse DNA
414	19	1.6 204407	2	AC021446	AC021446 Mus muscu	C 487	19	1.6 233109	2	AC116289	AC116289 Rattus no
415	19	1.6 205363	2	AC078998	AC078998 Mus muscu	C 488	19	1.6 233675	2	AC112444	AC112444 Rattus no
416	19	1.6 205677	2	AC068651	AC068651 Mus muscu	C 489	19	1.6 234983	2	AC099168	AC099168 Rattus no
417	19	1.6 206641	2	AC124321	AC124321 Mus muscu	C 490	19	1.6 236054	2	AC114393	AC114393 Rattus no
418	19	1.6 209527	2	AC115969	AC115969 Mus muscu	C 491	19	1.6 236475	3	AC094427	AC094427 Rattus no
419	19	1.6 209536	2	AC147213	AC147213 Pan trogl	C 492	19	1.6 236566	2	AC094427	AC094427 Rattus no
420	19	1.6 209972	2	AC134605	AC134605 Mus muscu	C 493	19	1.6 237400	2	AC118073	AC118073 Rattus no
421	19	1.6 210195	2	AC134605	AC134605 Mus muscu	C 494	19	1.6 237588	2	AC025581	AC025581 Mus muscu
422	19	1.6 211001	9	AC018695	AC018695 Homo sapi	C 495	19	1.6 238075	2	AC111727	AC111727 Rattus no
423	19	1.6 211173	9	AL671532	AL671532 Human DNA	C 496	19	1.6 238812	2	AC098077	AC098077 Rattus no
424	19	1.6 211395	2	AC138377	AC138377 Mus muscu	C 497	19	1.6 239298	2	AC116215	AC116215 Rattus no
425	19	1.6 211465	2	AC103070	AC103070 Rattus no	C 498	19	1.6 239434	2	AC073793	AC073793 Mus muscu
426	19	1.6 211873	10	AL596122	AL596122 Mouse DNA	C 499	19	1.6 240622	2	AC102528	AC102528 Mus muscu
427	19	1.6 213056	2	AC102350	AC102350 Mus muscu	C 500	19	1.6 242122	2	AC106448	AC106448 Rattus no
428	19	1.6 213163	2	AC123622	AC123622 Mus muscu	C 501	19	1.6 242499	2	AC119617	AC119617 Rattus no
429	19	1.6 213303	10	AL671520	AL671520 Mouse DNA	C 502	19	1.6 243674	3	AE003548	AE003548 Drosophila
430	19	1.6 213303	10	AL671520	AL671520 Mouse DNA	C 503	19	1.6 244309	2	AC107148	AC107148 Rattus no

504	19	1.6	244990	2	AC103233	AC103233 Rattus no	577	18	1.5	540	11	BV010417	BV010417 MASC STS1
505	19	1.6	245002	10	AC113276	AC113276 Mus muscu	c 578	18	1.5	545	11	BV010413	BV010413 MASC STS1
506	19	1.6	245257	2	AC135577	AC135577 Rattus no	579	18	1.5	554	11	BV010415	BV010415 MASC STS1
507	19	1.6	245312	2	AC113273	AC113273 Mus muscu	580	18	1.5	555	11	BV009599	BV009599 MASC STS1
508	19	1.6	245993	2	AC130970	AC130970 Rattus no	c 581	18	1.5	561	11	BV009601	BV009601 MASC STS1
509	19	1.6	246300	2	BX511023	BX511023 Danio rer	582	18	1.5	561	11	BV009600	BV009600 MASC STS1
510	19	1.6	246548	2	AC115131	AC115131 Rattus no	583	18	1.5	571	11	BV009598	BV009598 MASC STS1
511	19	1.6	247158	2	AC108807	AC108807 Mus muscu	584	18	1.5	576	11	BV009595	BV009595 MASC STS1
512	19	1.6	247176	2	AC097815	AC097815 Rattus no	585	18	1.5	584	11	BV059665	BV059665 S212P6159
513	19	1.6	248506	2	AC103085	AC103085 Rattus no	586	18	1.5	588	8	AY129472	AY129472 Arabidops
514	19	1.6	249509	2	AC130902	AC130902 Rattus no	587	18	1.5	597	6	AX576130	AX576130 Sequence
515	19	1.6	249509	2	AC130902	AC130902 Rattus no	588	18	1.5	667	9	AY448333	AY448333 Colobus g
516	19	1.6	249554	2	AC139417	AC139417 Rattus no	589	18	1.5	667	9	AY454885	AY454885 Colobus g
517	19	1.6	250552	2	AC109652	AC109652 Rattus no	590	18	1.5	687	6	AX867486	AX867486 Sequence
518	19	1.6	250969	2	AC111397	AC111397 Rattus no	591	18	1.5	687	6	BD147548	BD147548 Primer fo
519	19	1.6	251282	2	AC098155	AC098155 Rattus no	592	18	1.5	701	6	AX164245	AX164245 Sequence
520	19	1.6	251720	2	AC094479	AC094479 Rattus no	c 593	18	1.5	714	11	BV063407	BV063407 S212P6033
521	19	1.6	251957	2	AC126867	AC126867 Rattus no	c 594	18	1.5	725	6	BD011682	BD011682 Synthetic
522	19	1.6	252003	10	AL808105	AL808105 Mouse DNA	c 595	18	1.5	732	14	AF284644	AF284644 Sacbrood
523	19	1.6	254331	2	AC095285	AC095285 Rattus no	596	18	1.5	745	11	BV014108	BV014108 S212P6734
524	19	1.6	254958	2	AC095187	AC095187 Rattus no	c 597	18	1.5	772	8	ATH553542	ATH553542 Arabidops
525	19	1.6	257258	2	AC106905	AC106905 Rattus no	c 598	18	1.5	776	10	MMADH2807	MMADH2807 Mus musculu
526	19	1.6	258924	2	AC117845	AC117845 Rattus no	599	18	1.5	807	8	AY085620	AY085620 Arabidops
527	19	1.6	260168	2	AC094547	AC094547 Rattus no	599	18	1.5	824	8	AY039558	AY039558 Arabidops
528	19	1.6	260198	2	AC107136	AC107136 Rattus no	600	18	1.5	861	11	BV027597	BV027597 S212P6971
529	19	1.6	260665	2	AC113858	AC113858 Rattus no	601	18	1.5	892	9	AK000610	AK000610 Homo sapi
530	19	1.6	261493	2	AC106564	AC106564 Rattus no	602	18	1.5	914	6	BD184179	BD184179 Sequence
531	19	1.6	267349	2	AC134334	AC134334 Rattus no	603	18	1.5	914	6	BD192708	BD192708 123 human
532	19	1.6	269722	2	AC109043	AC109043 Rattus no	c 604	18	1.5	1021	11	G42856	G42856 Xq3656 KWOK
533	19	1.6	270287	2	AC141415	AC141415 Pan trogl	c 605	18	1.5	1035	8	AY124865	AY124865 Arabidops
534	19	1.6	272606	2	AC095738	AC095738 Rattus no	c 606	18	1.5	1111	4	P1GYSO2YM	P1GYSO2YM Sus scrofa
535	19	1.6	273918	2	AC127610	AC127610 Rattus no	c 607	18	1.5	1120	8	HASPAO	HASPAO Pichia anom
541	19	1.6	274026	2	AC129633	AC129633 Rattus no	c 608	18	1.5	1120	10	AF121979	AF121979 Mus muscu
536	19	1.6	275333	2	AC137410	AC137410 Rattus no	609	18	1.5	1138	5	AF370634	AF370634 Cyrtocara
537	19	1.6	275333	2	AC137410	AC137410 Rattus no	610	18	1.5	1139	5	AF370626	AF370626 Oxytaria
538	19	1.6	277968	2	AC109425	AC109425 Rattus no	c 611	18	1.5	1140	5	AF180875	AF180875 Barbus pe
539	19	1.6	280245	2	AC103319	AC103319 Rattus no	c 612	18	1.5	1141	5	AF287421	AF287421 Barbus by
540	19	1.6	283369	2	AC099121	AC099121 Rattus no	c 613	18	1.5	1141	5	AF287443	AF287443 Barbus pe
542	19	1.6	287328	2	AC112394	AC112394 Rattus no	c 614	18	1.5	1204	9	BC007966	BC007966 Homo sapi
543	19	1.6	287427	2	AC106914	AC106914 Rattus no	615	18	1.5	1211	8	BT004169	BT004169 Arabidops
544	19	1.6	299300	2	AC006881	AC006881 Caenorhab	616	18	1.5	1236	6	AR030967	AR030967 Sequence
545	19	1.6	300143	1	AE017165	AE017165 Prochloro	617	18	1.5	1239	8	AF367336	AF367336 Arabidops
546	19	1.6	340000	9	AP001707	AP001707 Homo sapi	618	18	1.5	1373	5	SACHORION	SACHORION Sparus aura
547	19	1.6	340000	9	HS21C083	HS21C083 Homo sapi	619	18	1.5	1380	6	AX595494	AX595494 Sequence
548	18	1.5	300	6	BD214006	BD214006 Novel hum	620	18	1.5	1380	6	AX819608	AX819608 Sequence
549	18	1.5	304	6	BD237202	BD237202 Metastati	621	18	1.5	1380	6	AX830638	AX830638 Sequence
550	18	1.5	304	6	AR240767	AR240767 Sequence	622	18	1.5	1386	9	HSU73737	HSU73737 Human X1-1/
551	18	1.5	314	14	HIM131647	HIM131647 Human imm	c 623	18	1.5	1409	3	AK114189	AK114189 Clona int
552	18	1.5	336	8	BT006516	BT006516 Arabidops	624	18	1.5	1430	10	AF237719	AF237719 Rattus no
553	18	1.5	346	6	BD060876	BD060876 Secreted	c 625	18	1.5	1443	6	AX300841	AX300841 Sequence
554	18	1.5	412	6	AR423864	AR423864 Sequence	626	18	1.5	1467	8	MTR311934	MTR311934 Medicago
555	18	1.5	412	6	BD119417	BD119417 EST and e	627	18	1.5	1506	8	AK070887	AK070887 Oryza sat
556	18	1.5	413	6	AR423862	AR423862 Sequence	c 628	18	1.5	1519	8	BT002399	BT002399 Arabidops
557	18	1.5	413	6	AR423870	AR423870 Sequence	c 629	18	1.5	1540	8	BT005713	BT005713 Arabidops
558	18	1.5	413	6	BD119415	BD119415 EST and e	630	18	1.5	1541	9	BC018788	BC018788 Homo sapi
559	18	1.5	414	6	BD119423	BD119423 EST and e	631	18	1.5	1578	9	AF178756	AF178756 Homo sapi
560	18	1.5	414	6	AR423867	AR423867 Sequence	c 632	18	1.5	1627	8	AY065074	AY065074 Arabidops
561	18	1.5	415	6	BD119420	BD119420 EST and e	c 633	18	1.5	1637	8	YSCPR3A	YSCPR3A Saccharomyc
562	18	1.5	415	6	AR423861	AR423861 Sequence	c 634	18	1.5	1670	8	AY088953	AY088953 Arabidops
563	18	1.5	415	6	BD119414	BD119414 EST and e	c 635	18	1.5	1676	9	HSM801277	HSM801277 Homo sapi
564	18	1.5	416	6	AR423863	AR423863 Sequence	636	18	1.5	1714	8	SCPROIS	SCPROIS S.cerevisia
565	18	1.5	416	6	AR423865	AR423865 Sequence	637	18	1.5	1774	8	RICAPPRO	RICAPPRO Oryza sativ
566	18	1.5	416	6	BD119416	BD119416 EST and e	638	18	1.5	1799	9	AF241831	AF241831 Homo sapi
567	18	1.5	416	6	BD119418	BD119418 EST and e	639	18	1.5	1842	6	AX405673	AX405673 Sequence
568	18	1.5	425	6	AR423869	AR423869 Sequence	c 640	18	1.5	1849	14	AK012548S2	AK012548S2 HIV-1 iso
569	18	1.5	425	6	BD119422	BD119422 EST and e	c 641	18	1.5	1900	8	AK065206	AK065206 Oryza sat
570	18	1.5	436	11	BV010414	BV010414 MASC STS1	c 642	18	1.5	2000	6	AX510586	AX510586 Sequence
571	18	1.5	442	10	AF039021	AF039021 Mus muscu	643	18	1.5	2028	9	AK025144	AK025144 Homo sapi
572	18	1.5	454	6	AR423868	AR423868 Sequence	c 644	18	1.5	2141	10	AF073309	AF073309 Mus muscu
573	18	1.5	454	6	BD119421	BD119421 EST and e	645	18	1.5	2188	10	BC061823	BC061823 Rattus no
574	18	1.5	513	8	AF280931S2	AF280931 Pongo pyg	646	18	1.5	2192	9	AK055161	AK055161 Homo sapi
575	18	1.5	523	8	AF084201	AF084201 Medicago	647	18	1.5	2207	6	AX401773	AX401773 Sequence
576	18	1.5	531	11	BV010407	BV010407 MASC STS1	648	18	1.5	2207	10	RNU10357	RNU10357 Rattus norv
577	18	1.5	534	11	BV010409	BV010409 MASC STS1	649	18	1.5	2216	5	BC054636	BC054636 Danio rer

650	18	1.5	2234	6	AX879644	AX879644 Sequence	C 723	18	1.5	10383	6	AX587586	AX587586 Sequence
651	18	1.5	2234	6	BD157947	BD157947 Primer fo	C 724	18	1.5	10388	9	HUMPOL1B	M73548 Human poly
652	18	1.5	2234	9	AX022706	AX022706 Homo sapi	C 725	18	1.5	10388	11	G28515	G28515 Human Sfs S
653	18	1.5	2240	6	AX598664	AX598664 Sequence	C 726	18	1.5	10646	8	AC015384	AC015384 Drosophil
654	18	1.5	2240	6	AX767366	AX767366 Sequence	C 727	18	1.5	10989	8	AF263518	AF263518 Arabidops
655	18	1.5	2240	6	AX795669	AX795669 Sequence	C 728	18	1.5	11047	1	AE012535	AE012535 Xanthomon
656	18	1.5	2273	8	SCYOR056C	274964 S. cerevisia	C 729	18	1.5	11237	1	AE006302	AE006302 Lactococc
657	18	1.5	2314	8	AX133830	AX133830 Arabidops	C 730	18	1.5	11361	1	AE000950	AE000950 Drosophil
658	18	1.5	2323	6	BD127627	BD127627 Primer fo	C 731	18	1.5	11422	3	AF190405	AF190405 Drosophil
659	18	1.5	2323	9	AX075271	AX075271 Homo sapi	C 732	18	1.5	11926	2	AC017527	AC017527 Drosophil
660	18	1.5	2348	5	BC063736	BC063736 Xenopus 1	C 733	18	1.5	12156	1	AE000759	AE000759 Aquifex a
661	18	1.5	2463	8	AX070037	AX070037 Arabidops	C 734	18	1.5	12395	4	SSU28757	SSU28757 Sus scrofa
662	18	1.5	2548	3	AE069861	AE069861 Drosophil	C 735	18	1.5	12975	1	AE004916	AE004916 Pseudomon
663	18	1.5	2558	3	CEY4486E	AL033511 Caenorhab	C 736	18	1.5	14789	1	AE000726	AE000726 Aquifex a
664	18	1.5	2666	3	AB062383	AB062383 Macaca ja	C 737	18	1.5	15400	1	AL591687	AL591687 Homo DNA
665	18	1.5	2687	3	PIIRTRANSC	L00947 Popillia ja	C 738	18	1.5	16697	9	AC138426	AC138426 Homo sapi
666	18	1.5	2688	3	AX095003	AX095003 Drosophil	C 739	18	1.5	23104	2	AC017874	AC017874 Drosophil
667	18	1.5	2704	6	AX699014	AX699014 Sequence	C 740	18	1.5	23936	9	AC109463	AC109463 Homo sapi
668	18	1.5	2715	9	F330116S06	AF330121 Homo sapi	C 741	18	1.5	25336	2	AC018183	AC018183 Drosophil
669	18	1.5	2942	8	S73876	S73876 FPR3-FKBP-7	C 742	18	1.5	24194	2	AC019463	AC019463 Drosophil
670	18	1.5	2991	6	AX147976	AX147976 Sequence	C 743	18	1.5	24339	6	AX349001	AX349001 Sequence
671	18	1.5	3067	9	AX091242	AX091242 Homo sapi	C 744	18	1.5	25335	8	SC8248	Z46373 S. cerevisia
672	18	1.5	3146	6	AX921106	AX921106 Sequence	C 745	18	1.5	25680	3	DRORYRH	D17389 Drosophila
673	18	1.5	3236	3	DMPKCR	M64362 Sweet potat	C 746	18	1.5	26591	1	AE008794	AE008794 Salmonell
674	18	1.5	3292	8	IPBSPP	IPBSPP	C 747	18	1.5	26591	1	AE008794	AE008794 Salmonell
675	18	1.5	3372	6	AX300832	AX300832 Sequence	C 748	18	1.5	26756	8	SPBC1198	AL355921 Schizosac
676	18	1.5	3373	9	AX017378	AX017378 Homo sapi	C 749	18	1.5	28924	8	AP004914	AP004914 Lotus cor
677	18	1.5	3377	9	AK126685	AK126685 Homo sapi	C 750	18	1.5	29523	3	AC024856	AC024856 Caenorhab
678	18	1.5	3380	9	AF095844	AF095844 Homo sapi	C 751	18	1.5	34484	2	AC015002	AC015002 Drosophil
679	18	1.5	3433	3	BT011070	BT011070 Drosophil	C 752	18	1.5	34509	3	CEFS7G12	CEFS7G12
680	18	1.5	3449	3	AF255741	AF255741 Drosophil	C 753	18	1.5	35467	9	AC005570	AC005570 Homo sapi
681	18	1.5	3840	6	AX463494	AX463494 Sequence	C 754	18	1.5	37221	2	AC013974	AC013974 Drosophil
682	18	1.5	4001	6	AX657831	AX657831 Sequence	C 755	18	1.5	37596	8	SPU33008	SPU33008 Schizosacch
683	18	1.5	4001	6	AX659105	AX659105 Sequence	C 756	18	1.5	40000	2	AR236806	AR236806 Sequence
684	18	1.5	4016	8	AK099567	AK099567 Oryza sat	C 757	18	1.5	40611	3	AC017952	AC017952 Drosophil
685	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 758	18	1.5	40611	3	AC017952	AC017952 Drosophil
686	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 759	18	1.5	41260	9	AC015339	AC015339 Drosophil
687	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 760	18	1.5	41260	9	AC015339	AC015339 Drosophil
688	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 761	18	1.5	41260	9	AC015339	AC015339 Drosophil
689	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 762	18	1.5	41260	9	AC015339	AC015339 Drosophil
690	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 763	18	1.5	41260	9	AC015339	AC015339 Drosophil
691	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 764	18	1.5	41260	9	AC015339	AC015339 Drosophil
692	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 765	18	1.5	41260	9	AC015339	AC015339 Drosophil
693	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 766	18	1.5	41260	9	AC015339	AC015339 Drosophil
694	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 767	18	1.5	41260	9	AC015339	AC015339 Drosophil
695	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 768	18	1.5	41260	9	AC015339	AC015339 Drosophil
696	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 769	18	1.5	41260	9	AC015339	AC015339 Drosophil
697	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 770	18	1.5	41260	9	AC015339	AC015339 Drosophil
698	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 771	18	1.5	41260	9	AC015339	AC015339 Drosophil
699	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 772	18	1.5	41260	9	AC015339	AC015339 Drosophil
700	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 773	18	1.5	41260	9	AC015339	AC015339 Drosophil
701	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 774	18	1.5	41260	9	AC015339	AC015339 Drosophil
702	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 775	18	1.5	41260	9	AC015339	AC015339 Drosophil
703	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 776	18	1.5	41260	9	AC015339	AC015339 Drosophil
704	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 777	18	1.5	41260	9	AC015339	AC015339 Drosophil
705	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 778	18	1.5	41260	9	AC015339	AC015339 Drosophil
706	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 779	18	1.5	41260	9	AC015339	AC015339 Drosophil
707	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 780	18	1.5	41260	9	AC015339	AC015339 Drosophil
708	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 781	18	1.5	41260	9	AC015339	AC015339 Drosophil
709	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 782	18	1.5	41260	9	AC015339	AC015339 Drosophil
710	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 783	18	1.5	41260	9	AC015339	AC015339 Drosophil
711	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 784	18	1.5	41260	9	AC015339	AC015339 Drosophil
712	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 785	18	1.5	41260	9	AC015339	AC015339 Drosophil
713	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 786	18	1.5	41260	9	AC015339	AC015339 Drosophil
714	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 787	18	1.5	41260	9	AC015339	AC015339 Drosophil
715	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 788	18	1.5	41260	9	AC015339	AC015339 Drosophil
716	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 789	18	1.5	41260	9	AC015339	AC015339 Drosophil
717	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 790	18	1.5	41260	9	AC015339	AC015339 Drosophil
718	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 791	18	1.5	41260	9	AC015339	AC015339 Drosophil
719	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 792	18	1.5	41260	9	AC015339	AC015339 Drosophil
720	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 793	18	1.5	41260	9	AC015339	AC015339 Drosophil
721	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 794	18	1.5	41260	9	AC015339	AC015339 Drosophil
722	18	1.5	4115	8	AK117122	AK117122 Arabidops	C 795	18	1.5	41260	9	AC015339	AC015339 Drosophil

796	18	1.5	63192	2	AC140595	AC140595 Macaca mu	869	18	1.5	97065	2	AP000768 Homo sapi
797	18	1.5	63227	9	AL357558	AL357558 Human DNA	870	18	1.5	98334	9	AC105038 Homo sapi
798	18	1.5	63478	10	AL929506	AL929506 Mouse DNA	871	18	1.5	99074	9	Z98751 Human DNA s
C 799	18	1.5	64082	2	AC010119	AC010119 Drosophila	C 872	18	1.5	99505	9	AP002087 Homo sapi
C 800	18	1.5	64840	2	AC137088	AC137088 Homo sapi	C 873	18	1.5	10000	9	AB020874 Homo sapi
C 801	18	1.5	65244	2	AC124302	AC124302 Homo sapi	C 874	18	1.5	10000	2	AC020372 Drosophila
C 802	18	1.5	65465	8	AB007650	AB007650 Arabidops	C 875	18	1.5	100452	9	AC004499 Homo sapi
C 803	18	1.5	65866	2	AC116462	AC116462 Homo sapi	C 876	18	1.5	100665	8	AC006533 Arabidops
C 804	18	1.5	67272	2	AC100076	AC100076 Mus muscu	C 877	18	1.5	101427	9	AL596269 Human DNA
C 805	18	1.5	67722	2	AC087747	AC087747 Homo sapi	C 878	18	1.5	101510	9	AL592524 Human DNA
C 806	18	1.5	68181	2	AC103972	AC103972 Homo sapi	C 879	18	1.5	101607	8	AP004334 Oryza sat
C 807	18	1.5	69260	2	AC017374	AC017374 Drosophila	C 880	18	1.5	101997	2	AP004238 Oryza sat
C 808	18	1.5	69434	9	AL731813	AL731813 Human DNA	C 881	18	1.5	102051	2	AC144345 Medicago
C 809	18	1.5	70387	3	AC087075	AC087075 Caenorhab	C 882	18	1.5	102271	8	AP006361 Lotus cor
C 810	18	1.5	70426	2	AC138371	AC138371 Homo sapi	C 883	18	1.5	102574	2	AC017935 Drosophila
C 811	18	1.5	71241	2	AC102061	AC102061 Mus muscu	C 884	18	1.5	102733	9	AP008747 Homo sapi
C 812	18	1.5	71310	2	AC068177	AC068177 Homo sapi	C 885	18	1.5	103023	9	AC011810 Arabidops
C 813	18	1.5	71347	2	AC025308	AC025308 Homo sapi	C 886	18	1.5	103157	8	AC011810 Arabidops
C 814	18	1.5	71971	2	AC013594	AC013594 Drosophila	C 887	18	1.5	103209	9	AL135798 Human DNA
C 815	18	1.5	72016	9	AC135075	AC135075 Homo sapi	C 888	18	1.5	103421	8	AP006142 Lotus cor
C 816	18	1.5	72393	2	AC101182	AC101182 Mus muscu	C 889	18	1.5	104389	9	AC090024 Homo sapi
C 817	18	1.5	72616	8	AP006091	AP006091 Lotus cor	C 890	18	1.5	104463	2	BX571814 Dario rer
C 818	18	1.5	72698	8	AB007645	AB007645 Arabidops	C 891	18	1.5	104534	9	AL356317 Human DNA
C 819	18	1.5	72902	8	ATAC015985	ATAC015985 Arabidops	C 892	18	1.5	104935	9	AC020942 Homo sapi
C 820	18	1.5	73935	2	AC011048	AC011048 Homo sapi	C 893	18	1.5	104995	9	AC026416 Homo sapi
C 821	18	1.5	74502	2	AC024300	AC024300 Homo sapi	C 894	18	1.5	105807	8	AC006085 Arabidops
C 822	18	1.5	76564	2	AC105120	AC105120 Homo sapi	C 895	18	1.5	105377	9	AC004132 Homo sapi
C 823	18	1.5	78181	8	AB011477	AB011477 Arabidops	C 896	18	1.5	106502	2	AL357142 Homo sapi
C 824	18	1.5	79054	8	F21F23	AC027656 Sequence	C 897	18	1.5	106839	9	HSJ07712
C 825	18	1.5	79546	2	AC139001	AC139001 Homo sapi	C 898	18	1.5	107057	9	HSJ07712
C 826	18	1.5	80167	8	AB022216	AB022216 Arabidops	C 899	18	1.5	107850	2	AC035148 Homo sapi
C 827	18	1.5	80446	8	T27D20	AF076274 Arabidops	C 900	18	1.5	108350	3	AC011761 Drosophila
C 828	18	1.5	80606	3	AC024843	AC024843 Caenorhab	C 901	18	1.5	108718	9	AC008837 Homo sapi
C 829	18	1.5	80992	9	AP005137	AP005137 Homo sapi	C 902	18	1.5	109359	9	AC004857 Homo sapi
C 830	18	1.5	81329	2	AC008300	AC008300 Drosophila	C 903	18	1.5	109491	2	AC022224 Homo sapi
C 831	18	1.5	82356	8	AC006135	AC006135 Arabidops	C 904	18	1.5	110000	2	AC031367 3
C 832	18	1.5	82393	3	AC003923	AC003923 Drosophila	C 905	18	1.5	110000	2	AC098448 1
C 833	18	1.5	83605	9	AB004220	AB004220 Homo sapi	C 906	18	1.5	110000	2	AC103177 2
C 834	18	1.5	83948	8	AB012246	AB012246 Arabidops	C 907	18	1.5	110000	2	AC103178 2
C 835	18	1.5	84023	8	AP005011	AP005011 Oryza sat	C 908	18	1.5	110000	2	AC107372 1
C 836	18	1.5	84748	2	AC116234 3	Continuation (4 of	C 909	18	1.5	110000	2	AC110355 2
C 837	18	1.5	84896	8	AB025639	AB025639 Arabidops	C 910	18	1.5	110000	2	AC119705 3
C 838	18	1.5	85210	2	AC011423	AC011423 Homo sapi	C 911	18	1.5	110000	2	AC119867 2
C 839	18	1.5	85408	2	AC120932 3	Continuation (4 of	C 912	18	1.5	110000	2	AC132558 2
C 840	18	1.5	85637	9	AL391002	AL391002 Human DNA	C 913	18	1.5	110000	2	BX470110 3
C 841	18	1.5	85837	8	SPU33010	U33010 Schizosacch	C 914	18	1.5	110000	2	LMFLCHR16_04
C 842	18	1.5	86027	9	AC137767	AC137767 Homo sapi	C 915	18	1.5	110000	2	LMFLCHR16_05
C 843	18	1.5	87256	9	AC091865	AC091865 Homo sapi	C 916	18	1.5	110000	2	LMFLCHR16_10
C 844	18	1.5	87501	5	EX001046	EX001046 Zebrafish	C 917	18	1.5	110000	2	PFMAL13_11
C 845	18	1.5	87925	10	AL929067	AL929067 Mouse DNA	C 918	18	1.5	110000	2	PFMAL13_12
C 846	18	1.5	88065	9	AP002015	AP002015 Homo sapi	C 919	18	1.5	110000	6	AR409405 1
C 847	18	1.5	88323	9	AL162423	AL162423 Human DNA	C 920	18	1.5	110000	6	BD061520 1
C 848	18	1.5	88343	8	AT119P19	AL022605 Arabidops	C 921	18	1.5	110000	9	AF440620 2
C 849	18	1.5	88548	8	ATF27H5	AL163852 Arabidops	C 922	18	1.5	110384	9	AC000117
C 850	18	1.5	88619	9	AC097723	AC097723 Homo sapi	C 923	18	1.5	111371	9	AF215841
C 851	18	1.5	88784	2	AC012997	AC012997 Drosophila	C 924	18	1.5	111373	9	AC103955 Homo sapi
C 852	18	1.5	88848	8	T24H24	AF075598 Arabidops	C 925	18	1.5	111389	9	AC093735 Homo sapi
C 853	18	1.5	89122	9	AC007199	AC007199 Homo sapi	C 926	18	1.5	111517	8	ATF11C18
C 854	18	1.5	89603	9	HSU63834	U33834 Human KIR p	C 927	18	1.5	111746	9	HS327J16
C 855	18	1.5	89626	9	AC004776	AC004776 Homo sapi	C 928	18	1.5	112268	8	AC093490
C 856	18	1.5	89826	8	ATF32A16	AL078468 Arabidops	C 929	18	1.5	112295	9	AC084337
C 857	18	1.5	93683	2	AC138685	AC138685 Homo sapi	C 930	18	1.5	112309	9	AC003025
C 858	18	1.5	94879	9	AC103965	AC103965 Homo sapi	C 931	18	1.5	113078	9	AC107209 Human Chr
C 859	18	1.5	95111	8	ATF27K19	AL163832 Arabidops	C 932	18	1.5	113703	5	AL831745 Zebrafish
C 860	18	1.5	95173	8	AC068324	AC068324 Arabidops	C 933	18	1.5	113799	9	AC012502
C 861	18	1.5	95364	2	AC079048	AC079048 Homo sapi	C 934	18	1.5	113975	2	AC117685
C 862	18	1.5	95417	9	AC010876	AC010876 Homo sapi	C 935	18	1.5	113975	2	AC117685
C 863	18	1.5	95442	9	AL161614	AL161614 Human DNA	C 936	18	1.5	115626	9	AC103792
C 864	18	1.5	95566	9	AL161614	AL161614 Human DNA	C 937	18	1.5	116174	10	AL691510 Mouse DNA
C 865	18	1.5	95670	8	AC026757	AC026757 Arabidops	C 938	18	1.5	116461	8	AC135232 Medicago
C 866	18	1.5	96080	9	AC097379	AC097379 Homo sapi	C 939	18	1.5	117157	8	AC034258 Oryza sat
C 867	18	1.5	96585	8	ATF10N7	AL021636 Arabidops	C 940	18	1.5	117720	8	AC137522 Medicago
C 868	18	1.5	96638	9	AC092545	AC092545 Homo sapi	C 941	18	1.5	118131	9	AC074132 Homo sapi







[illegible]



QY	625	GCTCGTCAAGAAATCACTTTCCCACTTGCTGCTTGTGATTCAGATGGTATGCTGCTGCTGCT	684
Db	601	GCTCGTCAAGAAATCACTTTCCCACTTGCTGCTTGTGATTCAGATGGTATGCTGCTGCTGCT	660
QY	685	GCTACTTTTTCTGTGCTATGGCATGTTTCATCGTTGGAGATTTAAAGTGATGACGACA	744
Db	661	GCTACTTTTTCTGTGCTATGGCATGTTTCATCGTTGGAGATTTAAAGTGATGACGACA	720
QY	745	GAAGCAAGAGATTCAGATGGAGGATCAGTGTCTTACTATGTCATTGATGATGATCACA	804
Db	721	GAAGCAAGAGATTCAGATGGAGGATCAGTGTCTTACTATGTCATTGATGATGATCACA	780
QY	805	GGAATAATATGCAAGGTTTCTTCTTAGAGCCATAGGATGTTGTTTGTGCATCATCA	864
Db	781	GGAATAATATGCAAGGTTTCTTCTTAGAGCCATAGGATGTTGTTTGTGCATCATCA	840
QY	865	CTAGCTTCTGCTGTTCTGATAGTGTCTGCTTCGGTGACTCAAGTCTTCGCGCTGCT	924
Db	841	CTAGCTTCTGCTGTTCTGATAGTGTCTGCTTCGGTGACTCAAGTCTTCGCGCTGCT	900
QY	925	TGTTTCCGGAGAGTTTCAGCAGAGAAAGTGTCTCTACTTCTTCTTCTTCTTCTTGGGA	984
Db	901	TGTTTCCGGAGAGTTTCAGCAGAGAAAGTGTCTCTACTTCTTCTTCTTCTTCTTGGGA	960
QY	985	TTTGTCTCTTACTTCTACGGCGAGTTTAAATCCGGCAGAAAGTGTGTGTAACCTCAA	1044
Db	961	TTTGTCTCTTACTTCTACGGCGAGTTTAAATCCGGCAGAAAGTGTGTGTAACCTCAA	1020
QY	1045	CCGCGGAGACAGAACTGCTATTCTTCC	1073
Db	1021	CCGCGGAGACAGAACTGCTATTCTTCC	1049
RESULT 5	AC021044	121720 bp DNA linear	PLN 30-OCT-2002
LOCUS	Arabidopsis thaliana chromosome I BAC F3H9 genomic sequence,		
DEFINITION	complete sequence.		
ACCESSION	AC021044	5	GI-8347959
VERSION	AC021044.5		
KEYWORDS	Arabidopsis thaliana (thale cress)		
SOURCE	Arabidopsis thaliana		
ORGANISM	Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta; Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi		
REFERENCE	1 (bases 1 to 121720)		
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Nguyen, M., Lam, B., Southwick, A., Miranda, M., Brooks, S., Buehler, E., Chao, Q., Chin, C., Chio, J., Choi, E., Gonzalez, A., How, B., Johnson-Hopson, C., Khan, S., Kim, C., Koo, T., Lee, J.M., Lenz, C., Liu, S., Mukhar, N., Pham, P., Sakano, H., Schwartz, J., Shinn, P., Thaveri, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J., and Davis, R.W.		
JOURNAL	Unpublished		
REFERENCE	2 (bases 1 to 121720)		
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Nguyen, M., Lam, B., Southwick, A., Bei, Q., Buehler, E., Chin, C., Chio, J., Choi, E., Dunn, P., Gonzalez, A., How, B., Kim, C., Koo, T., Lee, J.M., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukhar, N., Pham, P., Sakano, H., Schwartz, J., Shinn, P., Thaveri, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J., Theologis, A., and Davis, R.W.		
JOURNAL	Submitted (12-JAN-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA		
REFERENCE	3 (bases 1 to 121720)		
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Nguyen, M., Lam, B., Southwick, A., Bei, Q., Buehler, E., Chin, C., Chio, J., Choi, E., Dunn, P., Gonzalez, A., How, B., Kim, C., Koo, T., Lee, J.M., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukhar, N., Pham, P., Sakano, H., Schwartz, J., Shinn, P., Thaveri, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J., Theologis, A., and Davis, R.W.		
JOURNAL	Submitted (12-JAN-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA		
REFERENCE	4 (bases 1 to 121720)		
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Nguyen, M., Lam, B., Southwick, A., Bei, Q., Buehler, E., Chin, C., Chio, J., Choi, E., Dunn, P., Gonzalez, A., How, B., Kim, C., Koo, T., Lee, J.M., Lenz, C., Li, J., Liu, A., Liu, K., Liu, S., Mukhar, N., Pham, P., Sakano, H., Schwartz, J., Shinn, P., Thaveri, A., Toriumi, M., Vaysberg, M., Walker, M., Yu, G., Ecker, J., Theologis, A., and Davis, R.W.		
JOURNAL	Submitted (08-JUN-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA		
REFERENCE	5 (bases 1 to 121720)		
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Nguyen, M., Lam, B., Southwick, A., Ecker, J., Theologis, A., and Davis, R.W.		
JOURNAL	Submitted (11-AUG-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA		
REFERENCE	6 (bases 1 to 121720)		
AUTHORS	Federspiel, N.A., Palm, C.J., Conway, A.B., Conn, L., Hansen, N.F., Altafi, H., Nguyen, M., Lam, B., Southwick, A., Ecker, J., Theologis, A., and Davis, R.W.		
JOURNAL	Submitted (15-AUG-2000) DNA Sequencing and Technology Center, Stanford University, 855 California Avenue, Palo Alto, CA 94304, USA		
COMMENT	On Jun 8, 2000 this sequence version replaced gi:7340332. Bases 1-7,373 of clone F3H9 overlap with bases 96,792-104,163 of IGF BAC clone F3M18, gb AC010155 e-mail for correspondence: arabesequence.stanford.edu Genes with similarity to proteins in the databases are named 'putative', '-like' or 'similar to'. Genes that have EST similarity but no significant protein similarity are described as 'unknown proteins'. Genes that are annotated based only on gene prediction software are described as 'hypothetical proteins'. The gene prediction programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/section/index.html), GENSCAN (Chris Burge, http://gncmhc.stanford.edu/~chris/GENSCANW.html), and NetPlantGene (S.M. Hebe, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/NetPlantGene.html).		
FEATURES	Location/Qualifiers		
source	1. 121720		
	/organism="Arabidopsis thaliana"		
	/mol_type="genomic DNA"		
	/db_xref="taxon:3702"		
	/chromosome="I"		
	/clone="F3H9"		
misc_feature	1. 7373		
	/note="overlap with bases 96,792-104,163 of IGF BAC clone F3M18, gb AC010155, see GenBank record for clone F3M18 for annotation in this region."		
gene	9852..10816		
	/gene="F3H9.1"		
	/note="synonym: DRM1"		
CDS	join(9852..9957,10049..10274,10780..10816)		
	/gene="F3H9.1"		
	/note="Location of cDNA clone gb AF053746 and genomic clone gb AF053747."		
	/codon_start=1		
	/product="dormancy-associated protein"		
	/protein_id="AAF98422.1"		
	/db_xref="GI:9795604"		
	/translation="MWLEKLWDVAGPQDPGLGLRLKITTQTINIRDIGSSSK"		



<p>RESULT 6 PFSC04066 LOCUS DEFINITION Plasmodium falciparum DNA *** SEQUENCING IN PROGRESS *** from contig 4-66, complete sequence.</p>	<p>COMMENT Tel:81-438-52-3935, Fax:81-438-52-3934 Address for correspondence: kaos@kazu.or.jp For the latest information on annotation of this clone, please see http://www.kazu.or.jp/kaos/cgi-bin/aggd.graph.cgi?c=WMJ3 Genes with similarity to proteins in the databases are described in 'product' or 'note' qualifiers. Genes that have no significant protein similarity are described as 'unknown protein'. The software programs used to predict genes include: Grail (Informatics Group, Oak Ridge National Laboratory, http://compbio.ornl.gov/grail-1.3/), GENSCAN (Chris Burge, MIT, http://ccr-081.mit.edu/GENSCAN.html), NetGene2 (S.M. Hebsgaard, et al., CBS, Technical University of Denmark, http://www.cbs.dtu.dk/services/NetGene2/) and SplicePredictor (Volker Brendel, Stanford University, http://gremlin.zool.iastate.edu/cgi-bin/sp.cgi). Genes encoding tRNAs are predicted by tRNAscan-SE (Sean Eddy, Washington University School of Medicine, St. Louis, http://genome.wustl.edu/eddy/tRNAscan-SE/). This sequence may not be the entire insert of this clone. It may be shorter because we remove overlaps between neighboring submissions. The 5' clone is MF20 and the 3' clone is MYN21.</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 42356 /organism="Arabidopsis thaliana" /mol_type="genomic DNA" /strain="Columbia" /db_xref="taxon:3702" /chromosome="5" /clone="WMJ3" /clone_lib="Mitsui P1" join(3281..6554,6745..6840,6928..7382) /note="contains similarity to salt-inducible protein gene id:WMJ3.2" /codon_start=1 /evidence=not_experimental /protein_id="BA097283.1" /db_xref="GI:8843735" /translation="MCLTKDLNLLGSYETGDFMEKSYNTLTIDRWGSLNHHMDYROAR LRLVHGKALKFLKVVQKQPLETHIVQLVCIITHLVPMYDPAHLLKELSLMS GKSFVFGALMTVTRICNSPNSVVDILIRVLRGMIQDSLEIFRLMGLYGFNPSTVTKNALGVSQSGEDSVMSFKEMLRKKICPDVATFNILNVLCAEFGSPKSYLMOK MEKSGYAPTIYNTVTLHWYCKGRKAAITELDDHMKSGVDADVTYNNMLIHLDSHVNRIAGYLLRDMRKRMIHNEVNTYNTLINGSFNEGVLIASOLNMLFGLSPNHVTFNALIDHISSEGKFKALMFMVMEAGLTPPEVSQVLLDGLCKNAEPDLNGFCFKRMKRGVCGVRIITYGMDLCKNGFLDEAVLLNEMSGDIDPDITVYSALNGFCVGRFKTAKEIVRIYRVGLSPNGIYIYTLIYNCRCMGCKEAIKRIYEAMILEGTRDHPFNVLVTLCKAGKVAEAEFEMRCMTSDGILPNTVSDFCLINGYSGSGELKAFSVFDEMTKVGHPHTFTVGSLLKGLCKGHLREAEKFLSLHAPVAVDTVMYNTLLTAMC KSGNLAKAVSLFGEMVQSRILPDSYTVTSLISGLCRKGTVAIFLFAKEAARGNVLP NKVMYTCVDGMKAGQWQAGIYFREQMDNLGHTPDIVTNAMIDGYSRMGKIEKTNL LPEPQNQNGPNLTYYNILLHGYSKRQDVSTFLLYRSIILNGILPKLTCVSLVGL ICESNMLEIGLKILKAFICRGVEVDVRYTFNMLISKCCANGEINWAFDLVKVMTSLGIS LDKDTCDAMSVLNHRHQRQSRVLMHMSQGIQSPESRKYIGLGLCRVGDILKTAFLVKEEMIAHKICPPNVAESAMVRAKCGKADATLLRFLMKRLVPTIASFTPLMH LCCKGNVIEALELRVMSNGLKDLVSNVLTGLCAKGMALAFELVIEEMKDGFLANATYXALIRGLLARETAFSGADIILKDLARGFTMSLSQDSHNLKWMKELKALQNDIAVCHREELQSNVLSKNQLKFLCLHISRSLSLYCYTQDQKDSSEGTALAPFSDYKNSKTPRESLIESLKLPGGLINAVFNSTCIKDLVCLFVSYLLDSLAKSF FETKQEEEQALNSWSCVVISCAEFTCNQSGITKDNQSGKNDAPFQQKRASSVSLK ILDFYVGVTRN" join(8500..8505,9426..9614,9689..9733) /note="gene id:WMJ3.3" /codon_start=1 /evidence=not_experimental /product="NOI protein, nitrate-induced" /protein_id="BA097284.1" /db_xref="GI:8843736" /translation="MSDKGRPLPKFGEDWVNDPASAEFTVIFNKARDEKKTGGKPGS PKSSSEGHVKSQSGDPSKPPKPPKPKLCCMQAPAVDS" join(15029..15205,15277..17049) /note="gene id:WMJ3.4" /codon_start=1 /evidence=not_experimental</p>
<p>LOCUS PFSC04066 LOCUS DEFINITION Plasmodium falciparum DNA *** SEQUENCING IN PROGRESS *** from contig 4-66, complete sequence.</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>
<p>ACCESSION AL010241.1 GI:3273131</p>	<p>FEATURES 1. 5947 /organism="Plasmodium falciparum" /mol_type="genomic DNA" /db_xref="taxon:5833" /chromosome="4" /clone="4-66"</p>



/product="myosin heavy chain-like"  
/protein\_id="BAA97285.1"  
/db\_xref="GI:884373"  
/translation="MVAKGRRDSSSPIVEGEIDTSAPFQSVKDAVNLFGAFAFS  
AEKVFRRKPNQSAEKVLVQTEHLAQKLNKLEKNAETIREQALSELEWSKRT  
VDELTRKLEAVNSRDSANKATEAKSLIEAKPGNVSVASSDQATRDMEEGEVCK  
ELDTAKOELRKIQVSNILETETVALSKYVEAKKSVKSEKIELLRKIAAVNESV  
EQTKLACQARKQSEI/FAEKIQOKSKAGMEESAKSLAKNEFDFPEFAKLEVOL  
TEYNEIDELQOKETAKASDIDSVNGVSIENAKGLFEKLVBEESKALQELVESLKA  
ELKNVNEDEVAKEAIEISVAGDLHLKLSRSELEQCQVTEESKALQELVESLKA  
NOISSETAARAEAGNRKAKELMKEAENHLELSELHLRVALDEAEAKAAETK  
ALQIKMSKRTWAARNSSESQSIITLSOBEFKSLKRAEVDLFAEMKVAALA  
QVRAVASENETLKLTTQBEIKLKTATEEALKAAMADAAKAVEGELRWREED  
KGABENATRLAEAMKMASESPPQHIYKAPQKQPNVNNKLEKTKTSVSKVLMPLN  
QKAFNRKKNQVWGSYSLPGEKPF"  
complement (17520..18194)  
/notes="gene id:MMJ3.5"  
pi|T05237  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAA97286.1"  
/db\_xref="GI:8843738"  
/translation="MSLLNQPSKLLCFRKKVLRSSLMQTPPCSTVQVPCGADHL  
GRLDVMMSSRTCLEKVPLELVNSDPMVTIGSHGWATLKEDGILRLQDDLNPVAS  
DTATPKLPPLVTLPHCQTIQVLLQIKSTHCVFHLTYKQINTFFLTNTNQAHNLHF  
MDSKTLNSHPIKIKHQPSKYGLSVIYIINSINLLNQKKFHEIDFDMDFYILKNLFF  
CNLYFN"  
complement (18402..22877)  
/notes="gene id:MMJ3.6"  
/codon\_start=1  
/evidence=not experimental  
/product="retroelement pol polyprotein-like"  
/protein\_id="BAA97287.1"  
/db\_xref="GI:8843739"  
/translation="MSIAREVSSATPRTNQDPDVTKVSPYTLASSDNPGMATSSVWL  
TGNYNWSEMLNALQAKRTGFIINGSISKPLDNDPDYENWQAVNSMVGWIRASIE  
PKVSTVTFIDAHQWSELKQRFVGNKVRVHQIQAOLAACRQDQGPVIDYGRICK  
LWEEFOIYKTCGCLCTCGATLEPSEKREBEKIQHVLGLDSDRFGGLSATLIAM  
DPPFSGEIVSRVREERQIASVIREQQOQSAIGFLTRQSEVADGRDTSIIKSDR  
SVLSCGRCGRGHEKQCIQVGPDMWTERINGGRCSSRGGRSSGNNRGGRG  
QVTAANATNSLSPFEFTPDQRLVITOMIKNGTSDKLGKMKLGDVILDTGASH  
HMTGQSLNLTNITVIFSCSVGADRRKTPAISMTGFKLSTVSLNSVLYVPAKLSLI  
SVSKVQIKLALFTTICVLQDRFSRTLIGBERDGYIYLDATTTVHVVDVTT  
DHMLHQRGVPFVSLSPFLSGSCSVSSRCDVCFRAKQREVPFDPSSNKSTDC  
FSLHCDVWGPYRVPSCGAYFELTVDDPSRSVYTYLLAKSEVRSLTNFLAYTEK  
QFGSVKIIRSDNGTEMCLSSYFKEQIIVHOTSVCYCTPOONGRVERKRRHILNVSRA  
LRFQASLPKFGCEAVMTAYLINRTPSSIHNGLSPEYLLHGCKPDYDQLRVGSACY  
AHRVTRDKFGRSRLCIFVGFPGQWGVYDLSTNEFIVSRDVVRENVPFYATN  
EGDTIYTPVTCPTIDEDWLPFTLEDGRSDNSLSDPPVCTVDTSSESHTDTPQS  
LPTFVDPLSPSTVTPQTPTNSSSTSPSTNVSPQDTPPIENTPPQKRVQV  
QLARLQDYILYNASCTPTNTPLSPSTSSQSSSIQGNQSYPLTDYIFDECFSAHQVF  
LAATANDPEPKPEAKVKVMDAMVKEVDALEAVNTWDIVLPTGKVAIGSQWYK  
TKENAGTVERYKARLVQNNQIEGDIYETAPVVMVTITLRLVANQWQEVYQ  
MDVHNPAGDLGEVVKMLPPGRSHPKDVKCRKSLYGLKQACRNCFKLSLDALK  
RFQFIQGYEDYSPFSYCKGIELRVLYVDLLICGNDEYMQVKEFYLGRCSMDKL  
GKLYFLGIEVSGPGCIFLSQRYALDIISDSGTLCARPAYTPLEONHHLSDGDLPL  
LODPKPRRLVGLLYLHTRPESLVSVHVSOFMOAPRAHLEAMRIVRYLKGSPG  
QGILLNSKDLTLEVCDSPFQSCPLTRSLRAYVLLGSPISWTKKQDQTVSHSA  
EASYRWSVALKEIVKMLKLEIGTILAPTLFCDSKAAISIANPVFHERTKHIE  
RDCSHVRAVDGDIITTHVTRTSEQLADIFTKALGRNQFIYLMKSLGILQNLTPTT"  
complement (23987..25087)  
/notes="gene id:MMJ3.7"  
pi|T05237  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAA97288.1"  
/db\_xref="GI:8843740"  
/translation="MSLLNQPSKLLCFRKKVLRSSPLLSNGLSSSLRQTPTPTII  
DAVPCGEDGLKVIYANAVHFTYLEKKVLELVNDPMVTIGSHGWATLSQDDGIL  
RLQDDLNPASDTPNPKRIPLPLVTLPHCQTIQVTVNMSSSSPEDNCVAVKFLGP  
QISFCFPAQRNSWINVRIANPCFYSRVNMFSEKHDLFRIPGSGGHHIAGSWDLHKDH

CDS

CDS

CDS  
KNPKFQRLRFQNLPLKLTTKREVMDSYKSEHLVESRTTGETFLVKYKRVKAVRGMGSK  
LSTKALMVFRLDDEGNNAVTKDGLDCLIFLKSSEPPFCVLSLPRMFPNNVEYMDAD  
EDGYFNLRSSIVGDLTRMGTVGIYIPQNIIDN"  
complement (27620..28720)  
/notes="gene id:MMJ3.8"  
pi|T05237  
similar to unknown protein"  
/codon\_start=1  
/evidence=not experimental  
/protein\_id="BAA97289.1"  
/db\_xref="GI:8843741"  
/translation="MSLLNQPSKLLCFRKKVLRSSPLLSNGLSSSLRQTPTPTII  
DAVPCGEDGLKVIYANAVHFTYLEKKVLELVNDPMVTIGSHGWATLSQDDGIL  
RLQDDLNPASDTPNPKRIPLPLVTLPHCQTIQVTVNMSSSSPEDNCVAVKFLGP  
QISFCFPAQRNSWINVRIANPCFYSRVNMFSEKHDLFRIPGSGGHHIAGSWDLHKDH  
KNPKFQRLRFQNLPLKLTTKREVMDSYKSEHLVESRTTGETFLVKYKRVKAVRGMGSK  
LSTKALMVFRLDDEGNNAVTKDGLDCLIFLKSSEPPFCVLSLPRMFPNNVEYMDAD  
EDGYFNLRSSIVGDLTRMGTVGIYIPQNIIDN"  
complement (33024..36242)  
/notes="gene id:MMJ3.9"  
/codon\_start=1  
/evidence=not experimental  
/product="non-LTR retroelement reverse transcriptase-like"  
/protein\_id="BAA97290.1"  
/db\_xref="GI:8843742"  
/translation="MLGDFNQVLLPQEHNSPPSLNIDRRMRDFGSCISEMELSDLVFK  
GNSFTWNKSSIRPIAKDLRIILANDSWCNLYPSHGLFGNLDPSHDVSGVLEANG  
ISAKRPKPFNLLKNEDELNVMDNWFSTNVVSGSMYRVSKLAKMKPKIDKFSRLN  
YSGIELRTKEAHELLITCQNLTLANFSVSNAALELEAQRKVLSSCAESFFHQRSRV

CDS

Query Match 1.9%; Score 23; DB 8; Length 42356;  
Best Local Similarity 100.0%; Pred. No. 0.81;  
Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1172 TTTTGTATTGTTCTGTGTATA 1194  
|||||  
Db 11656 TTTTGTATTGTTCTGTGTATA 11634  
RESULT 8  
PFMAL4Pl/c  
LOCUS PFMAL4Pl 347582 bp DNA linear INV 29-JAN-2003  
DEFINITION Plasmodium falciparum MAL4Pl.  
ACCESSION AL034557 AL844503  
VERSION AL034557.8 GI:23498126  
KEYWORDS Plasmodium falciparum 3D7  
SOURCE Plasmodium falciparum 3D7  
ORGANISM Plasmodium falciparum 3D7  
Eukaryota; Alveolata; Apicomplexa; Haemosporida; Plasmodium.  
REFERENCE 1  
AUTHORS Hall, N., Pain, A., Berriman, M., Churcher, C., Harris, B., Harris, D.,  
Mungall, K., Bowman, S., Atkin, R., Baker, S., Barron, A., Brooks, K.,  
Buckee, C.O., Burrows, C., Cherevach, I., Chillingworth, C.,  
Chillingworth, T., Christodoulou, Z., Clark, L., Clark, R., Corton, C.,  
Cronin, A., Davies, R., Davis, P., Dear, P., Dearden, F., Doggett, J.,  
Feltwell, T., Goble, A., Goodhead, I., Gwilliam, R., Hamlin, N.,  
Hance, Z., Harper, D., Hauser, H., Hornsby, T., Holroyd, S.,  
Horrocks, P., Humphray, S., Jagels, K., James, K.D., Johnson, D.,  
Kerhornou, A., Knights, A., Konfortov, B., Kyes, S., Larke, N.,  
Lawson, D., Leonard, N., Line, A., Maddison, M., McLean, J., Mooney, P.,  
Moule, S., Murphy, L., Oliver, K., Ormond, D., Price, C., Quail, M.A.,  
Rabinowitz, E., Rajandream, M.A., Rutter, S., Rutherford, K.M.,  
Siddons, M., Simmonds, M., Seeger, K., Sharp, S., Smith, R., Squares, R.,  
Squares, S., Stevens, K., Taylor, K., Tivey, A., Unwin, L.,  
Whithead, S., Woodward, J., Sulston, J.E., Craig, A., Newbold, C. and  
Barrell, B.G.  
TITLE Sequence of Plasmodium falciparum chromosomes 1, 3-9 and 13  
JOURNAL Nature 419 (6906), 527-531 (2002)  
MEDLINE 22255708  
PUBMED 12368867

REFERENCE

AUTHORS

TITLE

JOURNAL

MEDLINE

PUBMED

REFERENCE

AUTHORS

CDS



<b>TITLE</b> JOURNAL	Direct Submission Submitted (24-FEB-1999) P.falciiparum Genome Sequencing Consortium, The Sanger Centre, Wellcome Trust Genome Campus, Hinxton, Cambridge CB10 1SA, UK
<b>COMMENT</b>	On Oct 3, 2002 this sequence version replaced gi:5731897. For more information about this sequence or the Malaria Project, see <a href="http://www.sanger.ac.uk/Projects/P_falciiparum">http://www.sanger.ac.uk/Projects/P_falciiparum</a> .
<b>FEATURES</b>	Location/Qualifiers
source	1. .347582 /organism="Plasmodium falciparum 3D7" /mol_type="genomic DNA" /isolate="3D7" /db_xref="taxon:36329" /chromosome="4" /clone="MAL4P1"
misc_feature	10384..32958 /note="rep20 Repeats"
gene	join(35153..41725,42757..44124) /gene="VAR"
CDS	/note="synonym: PFD0005w" join(35153..41725,42757..44124) /gene="VAR" /note="Similar to Plasmodium falciparum variant-specific surface protein var-2 SMALL:Q26033 (EMBL:I40609) (2664 aa) fasta scores: E(): 7.8e-149, 43.1% id in 2872 aa" /codon_start=1 /product="erythrocyte membrane protein 1 (PfEMP1)" /protein_id="CAD49094.1" /db_xref="GI:23498122" /db_xref="GOA:Q81220" /db_xref="SPTREMBL:Q81220" /translation="MVRIGSGGGVDDGDIDHQAHLDSIGKKHVDQVNGADGT GASGDAAKIDDLGDLKAPINPKLIGTDDPKLVEDYNNHNGDKGEYPCTE LSGKKQNPSPDLGGCTNSKMSKCEGACAPYRRLHLCNHNLESIEITKTASDTL LLEVCMAKTEKUEINHTYKHSNKDPSQLCTVLARSPADIGDIVRGDLFYGNT YESARKEKLEKUEVFGKLHGLSEAKKYGDDGNYQLREDWWTANRETVKAI TCEVSGNNYFRATCDGRNPSTLSQCRDCKDQKAGKPIKSGNNVIVPTVYVPOY LRWFEEAEDFCRLRKHKLDAIKKGRNKEBEYCDLNRYDCKNTASGHVFPFEDF CKQYCAPFDWIDQKLEFLKRYKTYKTEITSGSGCGSGRKKRDATTINYEYE KKYKELGTYKVNNFLKLENDVCTKNNDIKDGNIDFNHSGSAGKGGNNK TFYTKYCEAPGCAKVEGKAEKENCSTQKDYDPDKTTTIELTGDTKSDMVQ KYYKPCNGNGEKSATPNATSRKGGKGDOMEKVICYVDENKKYGSDAINFCVL QDGQHRKEQVTSYNAPFWKHVHMLHDSVRERLNSCINNASKNNKNEKNKE EKGJDAEAEKIEHLRQMLEQAGVRDLAAVGGPCTCGVAGETLAQVLAFLKLELDELKEAE CKNCPPKQAQEGPGARSADSPAGTEDHPDAEDDDDDDDDDDDDDDDDDDDDD PQCKVNDILSTDRTQVGDCHENKNGKPDWQCDLTLVDTTKVCPRRKLCL YIIAHSETKNIEQTDLRDAFRTAAETFLSWQYKIKNGADAKQLDNGTIPEEFL RSMFTYGDYHDCILNTDLSKTYNDVAKDKIGKPFSGKSPSGITTPODWQTY GKDIWGMKICALTHGVNTEKTKIKNDYSYKVNQSONGNSLSLEPAKKPOFLRWMI EWGEFCAERKLEQNTGKSCNGINPIQYCSDNRHPCNKACDIKYNVETKKEPRGQ TKFVRANLANDAQEDKYNTQPGSKQNDYLKDKCNKCSCEGNNVLTIVSSKK PFGIVAHKYSKCMGLAKFVTPNVPAPPQPPPPAIPAPATTPGVNPECIIVNTLF SDTKFWDACITKYGKAPTQWKICPTGNTSNEGAATDSEGSADKSRHRDLAPSSGS NQGSICVPPRRKLYVPLTKWASETTKGSQSQSGKAEGTESSESSSEASSPGTSSQ GERSQGLSTPASTSSNVDSDDLKAFVESAVETFLWHYKMDKOKELDEKKQ QRESGLVGALDNGSNVDDEDKQKLEGGDIPEEFKQMFVTLGDYDILVRGGNT SDSGNTSGNNNTVIEASGDQDEMKKIKQAIIDEKTIKDVNEVYKFFPTTQITSGTKE SSLTRETLEKHAHPSIWEGMICALTYYENDEKTIKDVNEVYKFFPTTQITSGTKE KYNTVTKDSDADYKCSGVGEBCKIEDISNIGVFLADLPCGCGRECKYKWTIER KLAQIKYDCKVDSADYKCSGVGEBCKIEDISNIGVFLADLPCGCGRECKYKWTIER KLEFNEQSDTFHTHGDPCFSGFKIDCBNGKSKSDTGKQCDGTLTIDAEIAMISS TPDVMVYSDNTTFEGDILKCEGSGGNGVCTGCTDTAAQFLERLPGCKNNDNGE KILKEEWEKIKKHYLEKNKEDNDMSVVRNLEKFEHRFEFNAIKRAQCNKCGVEM KLDQKQIILIRALIKRWLEYFLEDYNNKIKHISDCINNNGEINIKRAQCNKCGVEM FCLGNDKPSQNGHODADCMIKLEIDITSLSTSGEOTETECQHTPLDPDEDLL LEBENTPAPAEKKEKVEPOPOPPTTQIVKTPALVTSLAWSGVFGFAAFTYFFL EPPPTADILLKRYINPKSDVIDIPKLSNRYIPYTSKYGKRYIVLEGDSGTDGY KTKTSIDLLKRYINPKSDVIDIPKLSNRYIPYTSKYGKRYIVLEGDSGTDGY TDHSDITSSSSESEVEDINDIYAFRAPKTYKTLIEVLEPSSGNNTASGNNTASGN NTTASGNNTASGNNTASGNNTASGNNTASGNNTASGNNTASGNNTASGNNTASGN VNDYSSGDIPTFNTQHTNTLYDKPDEKPFITSIHDRNLITGEEYNDMSTNSGNDLY



This clone was selected for full length sequencing because it passed the following selection criteria: Similarity but not identity to protein.

```

FEATURES             source
  source
    Location/Qualifiers
      1..1996
        /organism="Danio rerio"
        /mol_type="mRNA"
        /strain="AB"
        /db_xref="taxon:7955"
        /clone="MGC:56719 IMAGE:2601210"
        /tissue_type="whole body, adult, (one male and one female, including unfertilized eggs)"
        /clone_lib="Sugano Kawakami zebrafish DRA"
        /lab_host="DH10B"
        /note="Vector: pME18S-FL3"
      100..720
        /codon_start=1
        /product="S-milar to RIKEN cDNA 0610011F06 gene"
        /protein_id="AAH45517.1"
        /db_xref="GI:28278022"
        /translation="MLNRAADRNKDPILSVLSRVASNRRLFALEISSGTGQHVHF
AKAFNITWQSEVEVTSLSSEAYRQVRLQVOPPIYLDVSQWOTWGGFPAESCD
LIINNMHISLDFLCTGLFHVGQGLKLPQGLLTYGPAENGSIQSNDFDQSLR
YRNPEWGLRDASFLLTLGQENGLREEIVDMPANNKCLLFRKDSVV"
      124..333
        /note="COG0220; Region: Predicted
S-adenosylmethionine-dependent methyltransferase [General
function prediction only]"
        /db_xref="CDD:COG0220"

misc_feature
  1.8%; Score 22; DB 5; Length 1996;
  Best Local Similarity 100.0%; Pred. No. 3.4;
  Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 ATTATGTGATGTTGTTT 1174
|||||
Db 1224 ATTATGTGATGTTGTTT 1245
|||||

RESULT 11
AY423021 2211 bp mRNA linear VRT 20-OCT-2003
LOCUS
DEFINITION
  Danio rerio hypothetical protein MGC13114 (MGC13114) mRNA, complete
  cds.
ACCESSION
  AY423021
VERSION
  AY423021.1 GI:37682140
KEYWORDS
  Danio rerio (zebrafish)
ORGANISM
  Danio rerio
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Actinopterygii; Neopterygii; Teleostei; Ostariophysi;
  Cypriniformes; Cyprinidae; Danio.
REFERENCE
  1 (bases 1 to 2211)
  Song,H.D., Wu,X.Y., Sun,X.J., Zhou,Y., Liu,T.X., Deng,M.,
  Zhang,G.W., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y.,
  Zou,L.I., Kanki,J.P., Look,A.T. and Chen,Z.
  Gene Expression Profiling in the Zebrafish Kidney Marrow Tissue
  Unpublished
  2 (bases 1 to 2211)
  Zhang,G.W., Sun,X.J., Wu,X.Y., Song,H.D., Zhou,Y., Liu,T.X.,
  Deng,M., Sheng,Y., Chen,Y., Ruan,Z., Jiang,C.L., Fan,H.Y.,
  Zou,L.I., Kanki,J.P., Look,A.T. and Chen,Z.
  Direct Submission
  Submitted (26-SEP-2003) State Key Lab for Medical Genomics,
  Shanghai Institute of Hematology, Ruijin Hospital Affiliated to
  Shanghai Second Medical University, 197 Rui Jin Road II, Shanghai
  200025, P. R. China
  Location/Qualifiers
    1..2211
      /organism="Danio rerio"
      /mol_type="mRNA"
      /db_xref="taxon:7955"

FEATURES             source
  source
    Location/Qualifiers
      1..2211
        /organism="Danio rerio"
        /mol_type="mRNA"
        /db_xref="taxon:7955"

```

```

/clone="RK067AlE08"
/tissue_type="kidney marrow"
1..2211
  /gene="MGC13114"
228..935
  /gene="MGC13114"
  /codon_start=1
  /product="hypothetical protein MGC13114"
  /protein_id="AAQ97997.1"
  /db_xref="GI:37682141"
  /translation="MNLSEGRHCPGTHLHRTSPKGTCTWCKMLNAAADRNKDPIL
SVLSRVASNRRLFALEISSGTGQHVHFARFAPFNITWQSEVEVTSLSSEAYRQVH
RLQVOPPIYLDVSQWOTWGGFPAESCDLIINNMHISLDFLCTGLFHVGQGLKLP
QGLLTYGPAENGSIQSNDFDQSLRYRNPEWGLRDASFLLTLGQENGLREEIV
DNPTNNKCLLFRKDSVV"

ORIGIN
  Query Match 1.8%; Score 22; DB 5; Length 2211;
  Best Local Similarity 100.0%; Pred. No. 3.3;
  Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 1153 ATTATGTGATGTTGTTT 1174
|||||
Db 1439 ATTATGTGATGTTGTTT 1460
|||||

RESULT 12
AC025719 44326 bp DNA linear HTG 12-MAR-2000
LOCUS
DEFINITION
  Caenorhabditis elegans clone Y45G8AM, *** SEQUENCING IN PROGRESS
  ***.
ACCESSION
  AC025719
VERSION
  AC025719.1 GI:7230343
KEYWORDS
  HTG; HTGS PHASE1.
SOURCE
  Caenorhabditis elegans
  ORGANISM
  Caenorhabditis elegans
  Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
  Rhabditidae; Rhabditidae; Peloderinae; Caenorhabditis.
  1 (bases 1 to 44326)
  Waterston,R.H.
  The sequence of Caenorhabditis elegans clone
  Unpublished
  2 (bases 1 to 44326)
  Waterston,R.H.
  Direct Submission
  Submitted (12-MAR-2000) Genome Sequencing Center, Washington
  University School of Medicine, 4444 Forest Park Parkway, St. Louis,
  MO 63108, USA
  * NOTE: This is a 'working draft' sequence. It currently
  * consists of 1 contigs. The true order of the pieces
  * is not known and their order in this sequence record is
  * arbitrary. Gaps between the contigs are represented as
  * runs of N, but the exact sizes of the gaps are unknown.
  * This record will be updated with the finished sequence
  * as soon as it is available and the accession number will
  * be preserved.
  * 1 44326: contig of 44326 bp in length.
  Location/Qualifiers
    1..44326
      /organism="Caenorhabditis elegans"
      /mol_type="genomic DNA"
      /db_xref="taxon:6239"
      /clone="Y45G8AM"

FEATURES             source
  source
    Location/Qualifiers
      1..44326
        /organism="Caenorhabditis elegans"
        /mol_type="genomic DNA"
        /db_xref="taxon:6239"
        /clone="Y45G8AM"

ORIGIN
  Query Match 1.8%; Score 22; DB 2; Length 44326;
  Best Local Similarity 100.0%; Pred. No. 3;
  Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 235 GAAACAAAGCGGAAACAAAGC 256
|||||
Db 29661 GAAACAAAGCGGAAACAAAGC 29682
|||||

```

RESULT 13	AC084155	46826 bp	DNA	linear	INV 04-OCT-2003
LOCUS	Caenorhabditis elegans cosmid Y45G5AM, complete sequence.				
DEFINITION	AC084155				
ACCESSION	AC084155.3	GI:16950482			
VERSION	HTG.				
KEYWORDS	Caenorhabditis elegans				
SOURCE	Caenorhabditis elegans				
ORGANISM	Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea; Rhabditidae; Pelodierinae; Caenorhabditis.				
REFERENCE	1 (bases 1 to 46826)				
AUTHORS	Wilson, R.				
TITLE	Genome sequence of the nematode C. elegans: a platform for investigating biology. The C. elegans Sequencing Consortium				
JOURNAL	Science 282 (5396), 2012-2018 (1998)				
MEDLINE	99069613				
PUBMED	9851916				
REFERENCE	2 (bases 1 to 46826)				
AUTHORS	Layman, D., Madsen, C. and Du, H.				
TITLE	The sequence of C. elegans cosmid Y45G5AM				
JOURNAL	Unpublished (2001)				
REFERENCE	3 (bases 1 to 46826)				
AUTHORS	Waterston, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (13-OCT-2000) Department of Genetics, Washington University, 4444 Forest Park Avenue, St. Louis, Missouri 63108, USA				
REFERENCE	4 (bases 1 to 46826)				
AUTHORS	Waterston, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (09-AUG-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA				
REFERENCE	5 (bases 1 to 46826)				
AUTHORS	Waterston, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-SEP-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA				
REFERENCE	6 (bases 1 to 46826)				
AUTHORS	Waterston, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (16-NOV-2001) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA				
REFERENCE	7 (bases 1 to 46826)				
AUTHORS	Waterston, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (24-MAY-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA				
REFERENCE	8 (bases 1 to 46826)				
AUTHORS	Waterston, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (07-JUN-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA				
REFERENCE	9 (bases 1 to 46826)				
AUTHORS	Waterston, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (06-NOV-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA				
REFERENCE	10 (bases 1 to 46826)				
AUTHORS	Waterston, R.				
TITLE	Direct Submission				
JOURNAL	Submitted (22-NOV-2002) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA				
REFERENCE	11 (bases 1 to 46826)				
AUTHORS	Waterston, R.				
TITLE	Direct Submission				

## JOURNAL

REFERENCE  
AUTHORS  
TITLE  
JOURNALREFERENCE  
AUTHORS  
TITLE  
JOURNAL

## COMMENT

FEATURES  
source

## gene

## CDS

Submitted (24-JAN-2003) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 12 (bases 1 to 46826)  
 Waterston, R.  
 Direct Submission  
 Submitted (07-APR-2003) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 13 (bases 1 to 46826)  
 Wilson, R.  
 Direct Submission  
 Submitted (04-OCT-2003) Department of Genetics, Washington University, Genome Sequencing Center, 4444 Forest Park Avenue, St. Louis, MO 63110, USA  
 On Nov 16, 2001 this sequence version replaced gi:15487595.  
 Submitted by: Genome Sequencing Center  
 Department of Genetics, Washington University  
 St. Louis, MO 63110, USA, and  
 Sanger Centre, Hinxton Hall  
 Cambridge CB10 1RQ, England  
 email: submissions@watson.wustl.edu and jessesanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once, or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one ml3 subclone.

For a graphical representation of this clone sequence and its analysis see:  
<http://www.wormbase.org/db/seq/sequence?name=Y45G5AM;class=Sequence>

## NEIGHBORING CLONE INFORMATION

The 5' clone is Y45G5AL, 3000 bp overlap; the 3' clone is Y45G5AR, 2300 bp overlap.

## NOTES:

Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program Genefinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yujii Kohara (http://www.ddbj.nig.ac.jp/c-elegans/html/CE\_INDEX.html) and The C. elegans ORFeome cloning project (http://worldb.dfci.harvard.edu/), similarity to other proteins from BlastX analyses (http://blast.wustl.edu/), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. RNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

## Location/Qualifiers

1. .46826  
 /organism="Caenorhabditis elegans"  
 /mol\_type="genomic DNA"  
 /strain="Bristol N2"  
 /db\_xref="taxon:6239"  
 /chromosome="V"  
 /clone="Y45G5AM"  
 complement(2833..7349)  
 /gene="Y45G5AM.6"  
 complement(join(2833..3009,3493..3611,4382..4491,4540..4668,5201..5444,5848..5928,6033..6111,7230..7349))  
 /gene="Y45G5AM.6"



Sanger Centre, Hinxton Hall  
Cambridge CB10 1RQ, England  
email: r.wenematode.wustl.edu and jee@sanger.ac.uk

NOTICE: This sequence may not be the entire insert of this clone. It may be shorter because we only sequence overlapping sections once or longer because we provide a small overlap between neighboring submissions.

This sequence was finished as follows unless otherwise noted: all regions were double stranded, sequenced with an alternate chemistry or covered by high quality data (i.e., phred quality >= 30); an attempt was made to resolve all sequencing problems, such as compressions and repeats; all regions were covered by sequence from more than one ml3 subclone.

For a graphical representation of this cosmid sequence and its analysis see:

<http://www.wormbase.org/db/seq/sequence?name=Y71F9B;class=Sequence>

#### NEIGHBORING COSMID INFORMATION

The 5' cosmid is F32B5, 600 bp overlap; the 3' cosmid is W03D8, 300 bp overlap.

#### NOTES:

Coding sequences below are the result of integration and manual review of the following data: computer analysis using the program Genefinder (P. Green and L. Hillier, personal communication), the large scale EST projects of Yuji Kohara ([http://www.dbbj.nig.ac.jp/c-elegans/html/CE\\_INDEX.html](http://www.dbbj.nig.ac.jp/c-elegans/html/CE_INDEX.html)) and The C. elegans ORFome cloning project (<http://wofdb.dfci.harvard.edu/>), similarity to other proteins from BlastX analyses (<http://blast.wustl.edu/>), sequence conservation with C. briggsae using Jim Kent's WABA alignment program (Genome Research 10:1115-1125, 2000), individual C. elegans GenBank submissions, and personal communications with C. elegans researchers. rRNAs are predicted using the program tRNAscan-SE (Lowe, T.M. and Eddy, S.R., 1997, Nucl. Acids. Res., 25, 955-964).

#### FEATURES

```

source
1..83297
/organism="Caenorhabditis elegans"
/mol_type="genomic DNA"
/strain="Bristol N2"
/db_xref="taxon:6239"
/chromosome="I"
/clone="Y71F9B"
6007..6008
/notes="Sl1 trans-splice acceptor; see yk513d5.5"
6038..15764
/genes="lin-17"
join(6038, 6103,6511..6615,8329..8475,9530..9814,
10848..10957,12260..12481,12529..12817,13137..13463,
15639..15764)
/genes="lin-17"
/standard_name="Y71F9B.5a"
/notes="contains similarity to Pfam domains PF01392 (Fz
domain), PF01534 (Frizzled/Smoothed family membrane
region); coded for by the following C. elegans cDNAs:
yk109c7.5, yk150f10.5, yk241b7.5, yk269h10.3, yk269h10.5,
yk275f4.3, yk275f4.5, yk341a1.3, yk341a1.5, yk496g4.3,
yk494d1.5, yk496g4.5, yk512h12.3, yk512h12.5, yk513d5.3,
yk622g4.3, yk622g4.5, yk648b3.3, yk648b3.5, U63557,
yk513d5.5"
/codon_start=1
/product="Abnormal cell lineage protein 17, isoform a"
/protein_id="AAF36028.1"
/db_xref="GI:7105632"
/db_xref="WormBase:Y71F9B.5a"
/translations="MHSLGIILLFPLATGSI FDOAVKGCIPIDIELCKLPYNT
YFNTILHNDQHTLQTHSHFPLMKTKCHPHIHFFICVFPAPMCIQFPAVTCKS
VCEQVADCFSLIEEFIGWPEPLNCAQFPDPPELCKPTEDEITGGFAPRLTKGS
SSSSKPTGCPSDLVDPDHPDKSHCAFCQSNMFSTDNKMRVRSWIWFAANAGV
ALFSLTFAIDRRKRPFRPCVYLSLIFISLPLYTLPLFIDAPIRSCHALNGRSY
VAGLSLATIIVLIFNKVDASELTALCSVGNLSIALFWFVIVPTICIVIGTCFIV
GFPASCRERISFRTRGTDTSKLEKVMKMGFFCALFILPNVIELVCCQYKFMILTOW
TRMTIDCKQSGCHRPFPQAEIYMTAVISSLATGFSCLMMWLSAKTWHAKNFIFC
GMSSAPVKNPIEPSTPLLEPPTAPPQPPVYTMQATNPQNSWRSPSKVY"
join(6038..6103,6511..6615,8329..8475,9530..9814,
10848..10957,12260..12481,12529..12817,13137..13467,
15639..15754)
/genes="lin-17"
/standard_name="Y71F9B.5b"
/notes="contains similarity to Pfam domains PF01392 (Fz
domain), PF01534 (Frizzled/Smoothed family membrane
region); coded for by the following C. elegans cDNAs:
yk275f4.3, yk648b3.3"
/codon_start=1
/product="Abnormal cell lineage protein 17, isoform b"
/protein_id="AAK82924.1"
/db_xref="GI:15055445"
/db_xref="WormBase:Y71F9B.5b"
/translations="MHSLGIILLFPLATGSI FDOAVKGCIPIDIELCKLPYNT
YFNTILHNDQHTLQTHSHFPLMKTKCHPHIHFFICVFPAPMCIQFPAVTCKS
VCEQVADCFSLIEEFIGWPEPLNCAQFPDPPELCKPTEDEITGGFAPRLTKGS
SSSSKPTGCPSDLVDPDHPDKSHCAFCQSNMFSTDNKMRVRSWIWFAANAGV
ALFSLTFAIDRRKRPFRPCVYLSLIFISLPLYTLPLFIDAPIRSCHALNGRSY
VAGLSLATIIVLIFNKVDASELTALCSVGNLSIALFWFVIVPTICIVIGTCFIV
GFPASCRERISFRTRGTDTSKLEKVMKMGFFCALFILPNVIELVCCQYKFMILTOW
TRMTIDCKQSGCHRPFPQAEIYMTAVISSLATGFSCLMMWLSAKTWHAKNFIFC
GMSSAPVKNPIEPSTPLLEPPTAPPQPPVYTMQATNPQNSWRSPSKVY"
complement(16537..19878)
/genes="dnj-30"
complement(join(16537..16738,17513..17676,19079..19399,
19849..19878))
/genes="dnj-30"
/standard_name="Y71F9B.16"
/notes="contains similarity to Pfam domain PF00226 (DnaJ,
prokaryotic heat shock protein); coded for by the
following C. elegans cDNAs: yk620a12.3, yk563e4.3,
yk114h5.5, yk563e4.5, yk620a12.5"
/codon_start=1
/product="DnaJ domain (prokaryotic heat shock protein)
protein 30"
/protein_id="AAK93866.1"
/db_xref="GI:15277401"
/db_xref="WormBase:Y71F9B.16"
/translations="MSQSFYQDLKSTEQQDAALTSEQOIERLLKPGSTYLNLPYE
VLQIDLDTIDEAKKKYKLSLVHPDKNDRADKAFIDIVKKAIKOMEDIELNR
CXDCYTEGARLAIVMSEKKRLKDKGKEDVIEDDDPAGYKALWITVTKVPADREKK

```

#### misc\_feature

```

Query Match      1.8%; Score 22; DB 3; Length 83297;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 235 GAAACAAAGCGGAAACAAAGC 256
      |||||
DB 72119 GAAACAAAGCGGAAACAAAGC 72140

RESULT 15
AC024854
LOCUS
Caenorhabditis elegans clone Y71F9B, complete sequence.
DEFINITION
AC024854
VERSION
AC024854.1 GI:7140417
KEYWORDS
HTG.
SOURCE
Caenorhabditis elegans
ORGANISM
Caenorhabditis elegans
Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida;
Rhabditidea; Rhabditidae; Peloderinae; Caenorhabditis.
REFERENCE
1 (bases 1 to 83297)
AUTHORS
Waterston, R.H.

```

Mon Sep 27 08:51:12 2004

```

TITLE      The sequence of Caenorhabditis elegans clone
JOURNAL    Unpublished
REFERENCE  2 (bases 1 to 83297)
AUTHORS    Waterston,R.H.
TITLE      Direct Submission
JOURNAL    Submitted (01-MAR-2000) Genome Sequencing Center, Washington
            University School of Medicine, 4444 Forest Park Parkway, St. Louis,
            MO 63108, USA
FEATURES   Location/Qualifiers
            source          1..83297
                           /organism="Caenorhabditis elegans"
                           /mol_type="genomic DNA"
                           /db_xref="taxon:6239"
                           /clone="Y71F9B"

ORIGIN
Query Match      1.8%; Score 22; DB 3; Length 83297;
Best Local Similarity 100.0%; Pred. No. 2.9;
Matches 22; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      235  GAAACACAGCGGAAACAAAGC 256
          |||||||
DB      72119  GAAACACAGCGGAAACAAAGC 72140

Search completed: September 25, 2004, 13:14:46
Job time : 5181 secs
```



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2004, 11:31:22 ; Search time 3627 Seconds  
(without alignments)  
10085.793 Million cell updates/sec

Title: US-09-913-767-1

Perfect score: 1225

Sequence: 1 aaacagcagcagcaagaa.....aatttgaagatattgagct 1225

Scoring table: OLIGO\_NUC  
Gapop 60.0 , Gapext 60.0

Searched: 27513289 seqs, 14931090276 residues

Word size : 0

Total number of hits satisfying chosen parameters: 55036578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database :

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_htc:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_gss\_hum:\*

18: em\_gss\_inv:\*

19: em\_gss\_pin:\*

20: em\_gss\_vrt:\*

21: em\_gss\_fun:\*

22: em\_gss\_mam:\*

23: em\_gss\_mus:\*

24: em\_gss\_pro:\*

25: em\_gss\_rod:\*

26: em\_gss\_pbg:\*

27: em\_gss\_vrl:\*

28: gb\_ges1:\*

29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	347	28.3	645	9	AU238506
2	242	19.8	438	9	AU229698
3	175	14.3	589	14	H76984
4	38	3.1	807	28	BZ501737
					AU238506 AU238506
					AU229698 AU229698
					H76984 17415 Lambd
					BZ501737 BONFA70TF

37	3.0	395	28	CC457536	CC457536 SALK_1106
31	2.5	784	28	BH601144	BH601144 BOHJF33TF
7	2.4	227	14	CD840768	CD840768 RFO2_120J
8	2.4	227	14	CD842095	CD842095 RFO2_12SP
26	2.1	613	28	BH657185	BH657185 BOMG919TF
26	2.1	728	28	BZ050387	BZ050387 Jnr42f08
11	2.1	761	28	BZ501745	BZ501745 BONFA70TR
26	2.1	824	28	BH601155	BH601155 BOHJF33TR
13	2.0	737	28	AQ901299	AQ901299 HS_5234_A
14	2.3	445	28	BH251397	BH251397 SALK_0115
23	1.9	1082	29	AG078331	AG078331 Pan_crog1
16	2.3	1189	28	CC294395	CC294395 CH261-2C1
17	2.2	597	10	BE129760	BE129760 945024B03
18	2.1	358	29	AG224068	AG224068 Lotus_cor
21	1.7	338	14	R37593	R37593 Yd09a06.s1
20	2.1	359	28	BH441465	BH441465 BOGO207TR
21	1.7	361	9	AI474435	AI474435 th20d10.x
21	1.7	367	14	R61853	R61853 YH10B03.s1
23	2.1	376	13	BY766870	BY766870 BY766870
24	1.7	475	9	AI342214	AI342214 qt38d05.x
25	2.1	495	13	BU494772	BU494772 PfESToab6
26	2.1	510	12	BG873698	BG873698 MEST8-F01
27	2.1	512	12	BG840197	BG840197 MEST8-F01
21	1.7	514	13	BQ407689	BQ407689 GA_Ed000
21	1.7	514	13	BQ407769	BQ407769 GA_Ed000
30	2.1	514	13	BQ407770	BQ407770 GA_Ed000
31	2.1	526	13	BQ407688	BQ407688 GA_Ed000
32	2.1	529	28	AZ725574	AZ725574 RPCI-24-1
33	2.1	538	14	C3625870	C3625870 wlin.pk01
34	2.1	539	10	AW936515	AW936515 QV4-DT002
35	2.1	548	13	CA142762	CA142762 SCMCRT210
36	2.1	556	14	CB605330	CB605330 3529_1_72
37	2.1	564	29	CG846730	CG846730 CMHD-GT_6
38	2.1	565	9	AA436973	AA436973 zv72h02.V
39	2.1	595	14	CD484641	CD484641 3529_1_11
40	2.1	688	12	BI405167	BI405167 MI-P-NA-a
41	2.1	707	13	BX504970	BX504970 DKFZp686J
42	2.1	767	28	BZ511825	BZ511825 BOMQ505TR
43	2.1	779	14	CD432518	CD432518 ETH1_30_C
44	2.1	808	10	B8054719	B8054719 GA_Ea003
45	2.1	827	28	BZ515292	BZ515292 BOMQL20TR
46	2.1	837	13	BU150941	BU150941 AGENCOURT
47	2.1	879	13	CA122327	CA122327 SCJLRL101
48	2.1	933	14	CF374942	CF374942 AGENCOURT
49	2.1	983	28	BZ485423	BZ485423 BONHS74TR
50	2.1	1033	28	BZ457602	BZ457602 BONHR74TR
51	2.1	1153	11	AY103575	AY103575 Zea_mays
52	2.0	103	10	BF937036	BF937036 EST460091
53	2.0	145	29	CG625403	CG625403 OST330949
54	2.0	160	12	BG736590	BG736590 rK59e08.Y
55	2.0	185	12	BI747102	BI747102 rK36a01.Y
56	2.0	185	12	BI747447	BI747447 rK40c07.Y
57	2.0	197	12	BG983420	BG983420 IL5-CN006
58	2.0	249	9	AV292764	AV292764 AV292764
59	2.0	257	29	CC934306	CC934306 ZMMB8C054
60	2.0	265	9	AI026602	AI026602 TENU0814
61	2.0	360	9	AV193425	AV193425 AV193425
62	2.0	360	13	C50658	C50658 C50658 Yui
63	2.0	360	13	C61160	C61160 C61160 Yui
64	2.0	360	13	C65289	C65289 C65289 Yui
65	2.0	375	13	C45990	C45990 C45990 Yui
66	2.0	383	28	B40799	B40799 HS-1052-B1
67	2.0	428	28	AQ581115	AQ581115 RPCI-11-4
68	2.0	439	28	BZ5564	BZ5564 F28C13TR_IG
69	2.0	455	14	CA613297	CA613297 wr1.pk015
70	2.0	465	10	BE468077	BE468077 hz70a08.x
71	2.0	476	9	AU288591	AU288591 AU288591
72	2.0	476	14	C3600217	C3600217 waw1c.pk0
73	2.0	507	12	BI174822	BI174822 OSTF054C6
74	2.0	510	29	C8593407	C8593407 tigr-gss-
75	2.0	525	10	AW560630	AW560630 EST15678
76	2.0	525	14	CD316132	CD316132 StrPu621.
77	2.0	555	14	CD316132	CD316132 StrPu621.



C 224	19	1.6	573	9	AV554688	AV554688	297	19	1.6	704	29	CE244424
C 225	19	1.6	573	14	CE270063	CE270063 1008970 H	298	19	1.6	706	14	AD869163
C 226	19	1.6	573	14	CF167955	B0792D09-H	299	19	1.6	708	9	AD612328
C 227	19	1.6	575	9	AI800839	AI800839 wgl3h04.x	300	19	1.6	710	14	CK265967
C 228	19	1.6	575	9	AI999672	AI999672 701556975	301	19	1.6	711	12	CK265967
C 229	19	1.6	577	9	AI498446	AI498446 tma49d04.x	302	19	1.6	711	29	CE078834
C 230	19	1.6	578	10	B4979993	B4979993 WHE0951.H	303	19	1.6	711	29	AG160893
C 231	19	1.6	580	12	B4979993	B4979993 WHE0951.H	304	19	1.6	712	29	AG160893
C 232	19	1.6	585	10	B4979993	B4979993 WHE0951.H	305	19	1.6	712	29	AG160893
C 233	19	1.6	588	14	B3555035	B3555035 B3555035	306	19	1.6	712	29	AG160893
C 234	19	1.6	588	14	CD467224	CD467224 LeukOS1.3	307	19	1.6	712	12	B1858456
C 235	19	1.6	588	28	BQ336854	BQ336854 HS 5018.B	308	19	1.6	714	12	B1858456
C 236	19	1.6	588	28	BH745745	BH745745 G77108.B	309	19	1.6	714	13	BY741089
C 237	19	1.6	593	13	B0609329	B0609329 BRY5254	310	19	1.6	715	12	BY741089
C 238	19	1.6	593	13	B0635577	B0635577 BRY5254	311	19	1.6	715	12	BY741089
C 239	19	1.6	594	9	AA042128	AA042128 004D01.In	312	19	1.6	716	12	BY741089
C 240	19	1.6	596	12	CBM307866	CBM307866 sak39a10.	313	19	1.6	716	12	BY741089
C 241	19	1.6	598	14	CD864874	CD864874 A202.001N	314	19	1.6	716	28	BH715556
C 242	19	1.6	600	12	BF1990106	BF1990106 4061-29.M	315	19	1.6	716	28	BH715556
C 243	19	1.6	601	12	CF169197	CF169197 B0810E03-	316	19	1.6	722	14	CF916975
C 244	19	1.6	601	12	BI462173	BI462173 603205175	317	19	1.6	722	14	CF916975
C 245	19	1.6	602	9	AV538672	AV538672 AV538672	318	19	1.6	727	28	BH933717
C 246	19	1.6	603	9	AV938036	AV938036 AV938036	319	19	1.6	730	12	BH933717
C 247	19	1.6	603	9	CB258152	CB258152 43-E01273	320	19	1.6	730	12	BH933717
C 248	19	1.6	605	12	B3533695	B3533695 B3533695	321	19	1.6	730	12	BH933717
C 249	19	1.6	605	12	CA608919	CA608919 wr1.PK009	322	19	1.6	737	12	BH933717
C 250	19	1.6	609	10	B336915	B336915 db68907.Y	323	19	1.6	738	9	AV117319
C 251	19	1.6	609	13	CA102166	CA102166 SCAGHR101	324	19	1.6	740	13	BQ443643
C 252	19	1.6	611	12	BI418857	BI418857 LJNEST36F	325	19	1.6	740	13	BQ443643
C 253	19	1.6	620	12	B3300929	B3300929 B3300929	326	19	1.6	743	14	B3532359
C 254	19	1.6	623	10	B499624	B499624 WHE0962.C	327	19	1.6	743	14	B3532359
C 255	19	1.6	624	13	BX620160	BX620160 BX620160	328	19	1.6	744	14	B3532359
C 256	19	1.6	627	29	CC947730	CC947730 BOIE742TR	329	19	1.6	744	14	B3532359
C 257												







C 808	18	1.5	585	13	BUT36928	881	18	1.5	617	14	CA597644
C 809	18	1.5	586	28	BZ948366	C 882	18	1.5	617	14	CF673979
C 810	18	1.5	585	14	CA302663	C 883	18	1.5	618	19	AV956090
C 811	18	1.5	586	28	AZ502328	C 884	18	1.5	618	29	CG166812
C 812	18	1.5	588	9	AV566910	C 885	18	1.5	619	14	CA755100
C 813	18	1.5	588	12	BM700135	C 886	18	1.5	620	29	CA755100
C 814	18	1.5	588	14	CF778060	C 887	18	1.5	622	13	CE054699
C 815	18	1.5	588	28	BZ294016	C 888	18	1.5	622	29	BU994998
C 816	18	1.5	589	14	CD056491	C 889	18	1.5	623	13	CG401307
C 817	18	1.5	590	14	AL786314	C 890	18	1.5	623	13	BU224453
C 818	18	1.5	590	14	CF009756	C 891	18	1.5	623	14	CB935688
C 819	18	1.5	590	29	CE049431	C 892	18	1.5	623	28	CD127473
C 820	18	1.5	591	13	BU980533	C 893	18	1.5	624	12	AZ421532
C 821	18	1.5	591	14	CF673995	C 894	18	1.5	624	13	BM70146
C 822	18	1.5	592	9	AI543681	C 895	18	1.5	625	14	BU006413
C 823	18	1.5	592	9	AV828053	C 896	18	1.5	625	14	CB888724
C 824	18	1.5	592	13	BU994742	C 897	18	1.5	625	29	CC948897
C 825	18	1.5	593	13	BU994064	C 898	18	1.5	627	12	BM098138
C 826	18	1.5	593	28	CC168535	C 899	18	1.5	627	13	BQ272626
C 827	18	1.5	594	9	AL775581	C 900	18	1.5	628	9	AL656826
C 828	18	1.5	594	10	AW201867	C 901	18	1.5	628	14	CD266488
C 829	18	1.5	594	14	CB637404	C 902	18	1.5	628	13	BM265169
C 830	18	1.5	595	14	CF798670	C 903	18	1.5	629	13	CE480902
C 831	18	1.5	595	14	CF798670	C 904	18	1.5	629	29	CE480902
C 832	18	1.5	596	14	CF675352	C 905	18	1.5	630	12	BG426540
C 833	18	1.5	596	28	BZ588552	C 906	18	1.5	631	14	CF600901
C 834	18	1.5	597	12	BG778770	C 907	18	1.5	631	14	CF600901
C 835	18	1.5	597	13	BQ660263	C 908	18	1.5	631	29	CC591146
C 836	18	1.5	597	14	CB935778	C 909	18	1.5	635	9	AJ453317
C 837	18	1.5	597	29	CE428265	C 910	18	1.5	635	10	AW694374
C 838	18	1.5	598	9	AL640028	C 911	18	1.5	635	10	BF299138
C 839	18	1.5	598	29	CG25547	C 912	18	1.5	635	13	BM082868
C 840	18	1.5	599	12	BG452496	C 913	18	1.5	637	10	BF643869
C 841	18	1.5	600	9	AV982976	C 914	18	1.5	638	13	BU733580
C 842	18	1.5	600	12	BG808458	C 915	18	1.5	638	28	BZ4361
C 843	18	1.5	600	28	AZ289836	C 916	18	1.5	638	29	CG450879
C 844	18	1.5	601	13	BQ368906	C 917	18	1.5	640	28	BZ941500
C 845	18	1.5	601	14	CF788266	C 918	18	1.5	641	13	CF654177
C 846	18	1.5	601	28	BH741292	C 919	18	1.5	641	13	BQ241548
C 847	18	1.5	601	28	BH741292	C 920	18	1.5	641	13	BQ241548
C 848	18	1.5	601	28	BZ718950	C 921	18	1.5	641	14	CA741593
C 849	18	1.5	601	28	BZ878678	C 922	18	1.5	641	14	CB918308
C 850	18	1.5	602	14	CF654671	C 923	18	1.5	641	28	BZ639733
C 851	18	1.5	603	9	AL700548	C 924	18	1.5	641	29	CG030709
C 852	18	1.5	603	12	BM580891	C 925	18	1.5	642	9	AI622723
C 853	18	1.5	603	13	CA137352	C 926	18	1.5	643	9	AI485483
C 854	18	1.5	603	29	CE752356	C 927	18	1.5	643	12	BI389664
C 855	18	1.5	604	14	CD394429	C 928	18	1.5	643	13	BU440284
C 856	18	1.5	604	14	W37319	C 929	18	1.5	643	14	CF654768
C 857	18	1.5	605	29	CE613208	C 930	18	1.5	644	28	CC145912
C 858	18	1.5	605	29	CE613208	C 931	18	1.5	644	9	AV821749
C 859	18	1.5	606	12	BG576291	C 932	18	1.5	644	14	CF674682
C 860	18	1.5	606	14	CB836571	C 933	18	1.5	645	29	CC728214
C 861	18	1.5	607	14	CA302441	C 934	18	1.5	646	12	BG584665
C 862	18	1.5	607	14	CA302441	C 935	18	1.5	646	29	CC747418
C 863	18	1.5	607	14	CF684810	C 936	18	1.5	647	12	BM112626
C 864	18	1.5	607	14	CF433757	C 937	18	1.5	647	29	AG252168
C 865	18	1.5	608	29	CE723876	C 938	18	1.5	648	12	BI419727
C 866	18	1.5	608	28	BZ791919	C 939	18	1.5	648	14	CF778940
C 867	18	1.5	608	28	BZ296707	C 940	18	1.5	648	29	CE314599
C 868	18	1.5	609	14	CB018333	C 941	18	1.5	650	9	AL849488
C 869	18	1.5	609	14	CF674498	C 942	18	1.5	650	14	CF913234
C 870	18	1.5	609	29	CE969834	C 943	18	1.5	650	28	CC106399
C 871	18	1.5	610	10	BF354457	C 944	18	1.5	650	29	CG655365
C 872	18	1.5	610	14	CF655742	C 945	18	1.5	651	9	AU236765
C 873	18	1.5	611	14	CB890897	C 946	18	1.5	651	9	AW067483
C 874	18	1.5	612	12	BJ371884	C 947	18	1.5	652	12	BG415556
C 875	18	1.5	612	14	CB271733	C 948	18	1.5	652	28	AF101588
C 876	18	1.5	613	9	AL677693	C 949	18	1.5	652	28	BZ110223
C 877	18	1.5	615	14	CB915342	C 950	18	1.5	652	29	CE071087
C 878	18	1.5	616	12	BJ041250	C 951	18	1.5	653	10	BB655772
C 879	18	1.5	616	28	BZ734804	C 952	18	1.5	653	12	BI957489
C 880	18	1.5	617	9	AI998902	C 953	18	1.5	654	10	BE203886
	18	1.5	617	13	CA015304		18	1.5	654	12	BG976551



954	18	1.5	654	12	BW15133	BW175133	UI-E-CL1-
C 955	18	1.5	654	13	BW045345	BW045345	30DGS--08
956	18	1.5	654	14	CF297669	CF297669	BP509859
957	18	1.5	656	12	BP509859	BP509859	taa47807.
C 958	18	1.5	656	14	CB889152	CB889152	taa61609.
C 959	18	1.5	655	14	CB889801	CB889801	ZMMBB5017
C 960	18	1.5	656	29	CR826938	CR826938	ZMMBB5017
C 961	18	1.5	657	12	BM665848	BM665845	UI-E-CL1-
C 962	18	1.5	657	13	BU007935	BU007935	QGH5P01.Y
C 963	18	1.5	657	28	AQ398889	AQ398889	mgxb00110
C 964	18	1.5	658	12	BG913316	BG913316	602812032
C 965	18	1.5	658	12	BJ532020	BJ532020	QJ532020
C 966	18	1.5	658	13	BQ850474	BQ850474	QGB12116.
C 967	18	1.5	659	14	CD268620	CD268620	taa96608.
C 968	18	1.5	659	14	CBW283678	CBW283678	BW283678
C 969	18	1.5	659	14	CB889757	CB889757	taa40H04.
C 970	18	1.5	659	29	CG917711	CG917711	MBEIK62TR
C 971	18	1.5	660	13	AQ482899	AQ482899	603849355
C 972	18	1.5	661	12	BJ541625	BJ541625	BJ541625
C 973	18	1.5	661	14	CD267519	CD267519	tab90C02.
C 974	18	1.5	662	13	BW206361	BW206361	BW206361
C 975	18	1.5	662	13	BI394550	BI394550	p9pin.pK0
C 976	18	1.5	663	12	BJ017231	BJ017231	BF017231
C 977	18	1.5	664	10	BP650606	BP650606	NF095B11E
C 978	18	1.5	664	14	CD566737	CD566737	tab43d08.
C 979	18	1.5	664	28	BZ831824	BZ831824	CH240.287
C 980	18	1.5	665	13	BW231239	BW231239	BW231239
C 981	18	1.5	666	13	BW098720	BW098720	BW098720
C 982	18	1.5	666	29	CE703004	CE703004	tigr-g88-
C 983	18	1.5	667	13	BY729536	BY729536	BY729536
C 984	18	1.5	667	14	CF674636	CF674636	tac94f08.
C 985	18	1.5	667	28	A6621751	A6621751	AC6271151
C 986	18	1.5	668	14	CF298319	CF298319	7LEAP--01
987	18	1.5	668	14	CF437041	CF437041	EST673386
988	18	1.5	668	14	CK023654	CK023654	AGENCOURT
989	18	1.5	669	12	B1196136	B1196136	602754619
990	18	1.5	669	13	BW229397	BW229397	BW229397
991	18	1.5	669	28	AZ836242	AZ836242	2M02131F05
992	18	1.5	669	29	CG347612	CG347612	OG0BM22TH
993	18	1.5	670	10	B8617656	B8617656	BB617656
994	18	1.5	670	12	BF170595	BF170595	BJ170595
995	18	1.5	670	14	CF777891	CF777891	tab21e12.
996	18	1.5	670	14	CF777937	CF777937	tab22C02.
997	18	1.5	670	14	CF780440	CF780440	tab220B04.
998	18	1.5	671	28	AZ952646	AZ952646	2M0217K06
999	18	1.5	675	12	BJ529315	BJ529315	BJ529315
999	18	1.5	675	14	CB672487	CB672487	OSJ5Ne061

RESULT 1	645 bp	linear	EST 01-APR-2002
AU238506			
LOCUS	645 bp	mRNA	
DEFINITION	AU238506 RAF117 Arabidopsis thaliana cDNA clone RAF117-32-017 5', mRNA sequence.		
ACCESSION	AU238506.1 GI:19877675		
VERSION	AU238506		
KEYWORDS	EST.		
SOURCE	Arabidopsis thaliana (thale cress)		
ORGANISM	Arabidopsis thaliana		
REFERENCE	1 (bases 1 to 645)		
AUTHORS	Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M., Akiyama, K., Enju, A., Ono, Y., Sakurai, T., Carninci, P., Kawai, J., Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shingawa, A., Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.		
TITLE	Large scale analysis of Arabidopsis full-length cDNA		
JOURNAL	Unpublished (2002)		
COMMENT	Contact: Motoaki Seki		

```

Db          629 CT 630

RESULT 2
AU229698/c
LOCUS      AU229698
DEFINITION AU229698 RAFL17 Arabidopsis thaliana cDNA clone RAFL17-32-017 3',
            mRNA sequence.
ACCESSION AU229698
VERSION   AU229698.1
KEYWORDS  AU229698.1 GI:19798174
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
           Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eutrosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 438)
AUTHORS   Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakejima,M.,
           Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
           Itoh,M., Ishii,Y., Arikawa,T., Shibata,K., Kende,H.,
           Muramatsu,M., Hayashizaki,Y. and Shinzaki,K.
TITLE     Large scale analysis of Arabidopsis full-length cDNA
JOURNAL   Unpublished (2002)
COMMENT   Contact: Motoaki Seki
           Plant Functional Genomics Research Group
           RIKEN Genomic Sciences Center
           3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
           Tel: 81-298-36-4359
           Fax: 81-298-36-9060
           Email: msek@rtc.riken.go.jp
           An Arabidopsis full-length cDNA library was constructed essentially
           as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
           and XhoI was ligated to modified Lambda FLC-1 vector (Carninci et
           al., submitted for publication) digested with BamHI and SalI. This
           clone is in a modified pluscript vector. Please visit our web
           site (http://www.sec.riken.go.jp/e/plant/index_e.html) for further
           details.

FEATURES             source
     source
     1..438
         /organism="Arabidopsis thaliana"
         /mol_type="mRNA"
         /db_xref="taxon:3702"
         /clone="RAFL17-32-017"
         /lab_host="DH10B"
         /clone_lib="RAFL17"
         /note="Site 1: BamHI; Site 2: SalI; Subtraction library.
           The sequence was obtained from samples subjected to
           dehydration-treated (1, 2, 5, 10 and 24 hr) and
           rehydration-treated (1, 2, 5, 10, and 24 hr after
           dehydration treatment)"

ORIGIN
Query Match      19.8%; Score 242; DB 9; Length 438;
Best Local Similarity 100.0%; Pred. No. 2.le-114;
Matches 242; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 819 AGCTTTCTTCTTAGGAGCCATAGGATGTTGTTGTCATCATCATAGCTTCTGGTGT 878
DB   |||
QY 879 TCTGATAAGTGTCTGCTTCGCGTGACTGAAGTTTTCGCGCGTGGTTGTTTCGGGAGAA 938
DB   |||
QY 939 GTTTCAGGAGAGAAAGGTGTCCTCTACTTTCTTTTCTTTGGGATTTGCTCTTACTT 998
DB   |||
QY 999 CTACGGCGAGTTAAATCCGGCAAGAAAGTTGTTGATAAACCTCAACCCGCGAGACAG 1058
DB   |||
QY 1058 CTACGGCGAGTTAAATCCGGCAAGAAAGTTGTTGATAAACCTCAACCCGCGAGACAG 1066
DB   |||

QY 1059 AC 1060
DB   ||
DB   165 AC 164

RESULT 3
H76984
LOCUS      H76984
DEFINITION H76984 Lambda-PRL2 Arabidopsis thaliana cDNA clone 200N2T7, mRNA
            sequence.
ACCESSION H76984
VERSION   H76984.1
KEYWORDS  H76984.1 GI:1054235
SOURCE    Arabidopsis thaliana (thale cress)
ORGANISM  Arabidopsis thaliana
           Arabidopsis thaliana
           Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
           Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
           rosids; eutrosids II; Brassicales; Brassicaceae; Arabidopsi.
REFERENCE 1 (bases 1 to 589)
AUTHORS   Newman,T., deBruijn,F.J., Green,P., Keegstra,K., Kende,H.,
           McIntosh,L., Ohlrogge,J., Raikhel,N., Somerville,S., Thomasow,M.,
           Retzel,E. and Somerville,C.
TITLE     Genes galore: a summary of methods for accessing results from
           large-scale partial sequencing of anonymous Arabidopsis cDNA clones
           Plant Physiol. 106, 1241-1255 (1994)
JOURNAL   95148729
MEDLINE   7846151
PUBMED    7846151
COMMENT   Contact: Thomas Newman
           MSU-DOE Plant Research Laboratory
           Michigan State University
           MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E.
           Lansing, MI
           Tel: 517-353-0854
           Fax: 517-353-9168
           Email: 22313tcm@ibm.cl.msu.edu
           Seq primer: T7 dye primer.
           Location/Qualifiers
           1..589
               /organism="Arabidopsis thaliana"
               /mol_type="mRNA"
               /strain="var Columbia"
               /db_xref="taxon:3702"
               /clone="200N2T7"
               /clone_lib="Lambda-PRL2"
               /note="Vector: Lambda Zip-Lox; Site 1: Sal; Site 2: Not;
                 Lambda PRL2 is a cDNA library derived from equal
                 quantities of 4 pools of mRNA. The mRNA sources were 1) 7
                 day germinated etiolated seedlings; 2) tissue culture
                 grown roots; 3) staged plants half with 24 hour light
                 cycle, half on 16 hr light, 8 hour dark- rosettes; 4)
                 same plants as 3 but aerial tissue (stems, flowers and
                 siliques. The vector is BRL's lambda Zip-Lox. The cDNA
                 inserts were directionally cloned with Sal-Not arms using
                 oligo dt primed cDNA."

ORIGIN
Query Match      14.3%; Score 175; DB 14; Length 589;
Best Local Similarity 99.8%; Pred. No. 2e-79;
Matches 225; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 55 ATTATCTCTCACTATAGAACATGTGGAGTCTTTGTTAACTCGTCTCTACTTCCACCAAT 114
DB   |||
QY 115 GCGGAAAAACGAATCTGGTTTCATGAGTTCCTATCAACCGCTGGTTTCAATCATCTCTC 174
DB   |||
QY 61 GCGGAAAAACGAATCTGGTTTCATGAGTTCCTATCAACCGCTGGTTTCAATCATCTCTC 120
DB   |||
QY 175 ATCCCTCTCTTGGTCTCTCTTCTTCCAGCGTCCGCGCGGCAACCGCAACCTTACACGCG 234
DB   |||
QY 121 ATCCCTCTCTTGGTCTCTCTTCTTCCAGCGTCCGCGCGGCAACCGCAACCTTACACGCG 180
DB   |||
QY 235 GAAAAACAGCGGAAAAACAAAGCTCTTCTCTCATGAAACTCTCTGT 280
DB   |||

```

Db 181 GAAACAGCGGAAACAAAGCTTCTCTCATGTGAACTCTCTGT 226

|||||

RESULT 4  
BZ501737/c  
LOCUS  
DEFINITION BZ501737 BO 1.6.2 KB tot Brassica oleracea genomic clone BONFA70,  
genomic survey sequence.  
ACCESSION BZ501737 807 bp DNA linear GSS 16-DEC-2002  
VERSION BZ501737.1 GI:27018757  
KEYWORDS  
SOURCE  
ORGANISM Brassica oleracea  
Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 807)  
AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
TITLE Whole genome shotgun sequencing of Brassica oleracea  
JOURNAL Unpublished (2001)  
COMMENT Other GSSs: BONFA70TR  
Contact: Chris Town

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.  
Location/Qualifiers  
1. .807  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BONFA70"  
/clone\_lib="BO 1.6.2 KB tot"  
/notes="Vector: PHOS1; Site 1: BatXI; 1.6-2 kb sheared  
total DNA inserted into PHOS1 using BatXI linkers"

FEATURES  
source

ORIGIN  
Query Match 3.1%; Score 38; DB 28; Length 807;  
Best Local Similarity 100.0%; Pred. No. 6e-08; Indels 0; Gaps 0;  
Matches 38; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 928 TTCGGGAGAGTTTCAGGCAGAGAAAGTGTCTCTCT 965  
Db 797 TTCGGGAGAGTTTCAGGCAGAGAAAGTGTCTCTCT 760  
|||||

RESULT 5  
CC457536  
LOCUS  
DEFINITION CC457536 395 bp DNA linear GSS 30-MAY-2003  
SALK\_110605.26.10.x Arabidopsis thaliana TDNA insertion lines  
Arabidopsis thaliana genomic clone SALK\_110605.26.10.x, genomic  
survey sequence.  
ACCESSION CC457536  
VERSION CC457536.1 GI:31219753  
KEYWORDS  
SOURCE  
ORGANISM Arabidopsis thaliana (thale cress)  
Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
REFERENCE 1 (bases 1 to 395)  
AUTHORS Alonso,J.M., Leisse,T.J., Barajas,P., Chen,H., Cheuk,R.,  
Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,  
Sinn,P., Zimmerman,J. and Ecker,J.R.  
TITLE A Sequence-Indexed Library of Insertion Mutations in the  
Arabidopsis Genome  
JOURNAL Unpublished (2001)

COMMENT Contact: Joseph R. Ecker  
Salk Institute Genomic Analysis Laboratory (SIGNAL)  
The Salk Institute for Biological Studies  
10010 N. Torrey Pines Road, La Jolla, CA 92037, USA  
Tel: 858 453 4100 x1752  
Fax: 858 558 6379  
Email: ecker@salk.edu  
This is single pass sequence recovered from the left border of  
TDNA. This sequence lies within an annotated exon of At1g28230.  
Class: TDNA tagged.  
Location/Qualifiers  
1. .395  
/organism="Arabidopsis thaliana"  
/mol\_type="genomic DNA"  
/strain="Columbia 0"  
/db\_xref="taxon:3702"  
/clone="SALK\_110605.26.10.x"  
/clone\_lib="Arabidopsis thaliana TDNA insertion lines"  
/notes="PCR was performed on Arabidopsis thaliana lines  
each of which contains one or more TDNA insertion  
elements. The resultant fragment for each line was  
directly sequenced to determine the genomic sequence at  
the site of insertion. Details of the protocols used can  
be found at http://signal.salk.edu/tdna\_protocols.html"

ORIGIN  
Query Match 3.0%; Score 37; DB 28; Length 395;  
Best Local Similarity 100.0%; Pred. No. 1.6e-07; Indels 0; Gaps 0;  
Matches 37; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1007 AGTTTAATCCGCGAGAAAGTTGTTGATAAACCTCA 1043  
Db 116 AGTTTAATCCGCGAGAAAGTTGTTGATAAACCTCA 152  
|||||

RESULT 6  
BH601144  
LOCUS  
DEFINITION BH601144 784 bp DNA linear GSS 15-DEC-2001  
survey sequence.  
ACCESSION BH601144  
VERSION BH601144.1 GI:17853590  
KEYWORDS  
SOURCE  
ORGANISM Brassica oleracea  
Brassica oleracea  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
REFERENCE 1 (bases 1 to 784)  
AUTHORS Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.  
TITLE Whole genome shotgun sequencing of Brassica oleracea  
JOURNAL Unpublished (2001)  
COMMENT Other GSSs: BOHJF33TR  
Contact: Chris Town

TIGR  
9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TF  
Class: sheared ends.  
Location/Qualifiers  
1. .784  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOHJF33"  
/clone\_lib="BOHJ"  
/notes="Vector: PHOS1; Site 1: BatXI; 2-3 kb sheared  
genomic DNA inserted into PHOS1 using BatXI linkers"

FEATURES  
source

ORIGIN

```
Query Match      2.5%; Score 31; DB 28; Length 784;
Best Local Similarity 100.0%; Pred. No. 0.00027;
Matches 31; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 65 CTATAGAACATCTGGAGGTCCTTTGTTAAAC 95
|||||
Db 204 CTATAGAACATCTGGAGGTCCTTTGTTAAAC 234

RESULT 7
CD840768
LOCUS
DEFINITION
RPO2.125P19F010627 RPO2 Brassica napus cDNA clone RPO212016, mRNA
sequence.
ACCESSION
CD840768
VERSION
CD840768.1 GI:32522708
KEYWORDS
EST.
SOURCE
Brassica napus (rape)
ORGANISM
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 227)
AUTHORS
Genoplante.
TITLE
Genoplante, a major partnership french program in plant genomics
JOURNAL
Unpublished (2003)
COMMENT
Contact: Genoplante
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).

FEATURES
source
Location/Qualifiers
1..227
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="samourai (restored line)"
/db_xref="taxon:3708"
/clone="RPO2125P19"
/tissue_type="anthers"
/clone_lib="RFO2"

ORIGIN
Query Match      2.4%; Score 29; DB 14; Length 227;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 991 TCTTACTTCTACGGCGAGTTTAATCCGG 1019
|||||
Db 22 TCTTACTTCTACGGCGAGTTTAATCCGG 50

RESULT 8
CD842095
LOCUS
DEFINITION
RPO2.125P19F010622 RPO2 Brassica napus cDNA clone RPO2125P19, mRNA
sequence.
ACCESSION
CD842095
VERSION
CD842095.1 GI:32524035
KEYWORDS
EST.
SOURCE
Brassica napus (rape)
ORGANISM
Brassica napus
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 227)
AUTHORS
Genoplante.
TITLE
Genoplante, a major partnership french program in plant genomics
JOURNAL
Unpublished (2003)
COMMENT
Contact: Genoplante
```

```
Genoplante
93, rue Henri Rochefort 91025 EVRY CEDEX France
Tel: 33 1 69 47 54 00
Fax: 33 1 69 47 54 10
This sequence has been generated in the framework of the french
plant genomics programme 'Genoplante' (http://www.genoplante.com
and http://genoplante-info.infobiogen.fr).

FEATURES
source
Location/Qualifiers
1..227
/organism="Brassica napus"
/mol_type="mRNA"
/cultivar="samourai (restored line)"
/db_xref="taxon:3708"
/clone="RFO2125P19"
/tissue_type="anthers"
/clone_lib="RFO2"

ORIGIN
Query Match      2.4%; Score 29; DB 14; Length 227;
Best Local Similarity 100.0%; Pred. No. 0.0019;
Matches 29; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 991 TCTTACTTCTACGGCGAGTTTAATCCGG 1019
|||||
Db 22 TCTTACTTCTACGGCGAGTTTAATCCGG 50

RESULT 9
BH657185
LOCUS
DEFINITION
BOMGP19TF BO 2.3 KB Brassica oleracea genomic clone BOMGP19,
genomic survey sequence.
ACCESSION
BH657185
VERSION
BH657185.1 GI:18715552
KEYWORDS
GSS.
SOURCE
Brassica oleracea
ORGANISM
Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
REFERENCE
1 (bases 1 to 613)
AUTHORS
Town, C.D., Van Aken, S., Utterback, T., Koo, H. and Fraser, C.M.
TITLE
Whole genome shotgun sequencing of Brassica oleracea
JOURNAL
Unpublished (2001)
COMMENT
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TF
Class: sheared ends.
Location/Qualifiers
1..613
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOMGP19"
/clone_lib="BO 2.3 KB"
/note="Vector: pHOS1; Site 1: BatXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

FEATURES
source
Location/Qualifiers
1..613
/organism="Brassica oleracea"
/mol_type="genomic DNA"
/strain="TO1000DH3"
/db_xref="taxon:3712"
/clone="BOMGP19"
/clone_lib="BO 2.3 KB"
/note="Vector: pHOS1; Site 1: BatXI; 2-3 kb sheared
genomic DNA inserted into pHOS1 using BstXI linkers"

ORIGIN
Query Match      2.1%; Score 26; DB 28; Length 613;
Best Local Similarity 100.0%; Pred. No. 0.099;
Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 118 GGAACAGCAATCTGGTTTCATGACGTT 143
|||||
Db 570 GGAACAGCAATCTGGTTTCATGACGTT 595
```

Mon Sep 27 08:51:13 2004

```

RESULT 10
BZ050387/c
LOCUS
DEFINITION
  jnr42f08.g1 B.oleracea001 Brassica oleracea genomic, genomic survey
  sequence.
  728 bp DNA linear GSS 09-OCT-2002
ACCESSION
BZ050387
VERSION
BZ050387.1 GI:23649836
KEYWORDS
GSS.
SOURCE
  Brassica oleracea
  Brassica oleracea
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
  1 (bases 1 to 728)
REFERENCE
  Delehaunty,K., Fewell,G., Fulton,L., McCombie,W.R., Miner,T.,
  Nash,W., Rabinowicz,P.D. and Wilson,R.K.
  Whole genome shotgun reads from Brassica oleracea
  Unpublished (2002)
  Contact: Richard K. Wilson
  Genome Sequencing Center
  Washington University School of Medicine
  Email: submissions@watson.wustl.edu
  Plate: jnr42 row: f column: 08
  Seq primer: -28RpPOT reverse
  Class: shotgun
  High quality sequence start: 33
  High quality sequence stop: 202.
  Location/Qualifiers
    1..728
      /organism="Brassica oleracea"
      /mol_type="genomic DNA"
      /db_xref="taxon:3712"
      /clone_lib="B. oleracea001"
      /notes="Vector: pOTw13; Whole genome shotgun library from
      flowering buds. DNA was purified from a crude nuclear
      prep using Brassica oleracea TO1000DH3 buds provided by
      Thomas Osborn at the University of Wisconsin. Genomic
      DNA was provided by Pablo Rabinowicz (CSHL) and the
      shotgun library prepared at Washington University Genome
      Sequencing Center."
FEATURES
  source
    1..728
      Query Match 2.1%; Score 26; DB 28; Length 728;
      Best Local Similarity 100.0%; Pred. No. 0.1;
      Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 979 TGGGGATTGCTCTTACTTCTACGG 1004
  |||||||||||||||||||||||||||
Db 126 TGGGGATTGCTCTTACTTCTACGG 101
  |||||||||||||||||||||||||||

RESULT 11
BZ501745
LOCUS
DEFINITION
  BONFA70TR BO_1.6_2_KB_tot Brassica oleracea genomic clone BONFA70,
  genomic survey sequence.
  761 bp DNA linear GSS 16-DEC-2002
ACCESSION
BZ501745
VERSION
BZ501745.1 GI:27018772
KEYWORDS
GSS.
SOURCE
  Brassica oleracea
  Brassica oleracea
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
  1 (bases 1 to 761)
REFERENCE
  Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  Whole genome shotgun sequencing of Brassica oleracea
  Unpublished (2001)
  Other_GSSs: BONFA70TF
  Contact: Chris Town
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA.
  Query Match 2.1%; Score 26; DB 28; Length 824;
  Best Local Similarity 100.0%; Pred. No. 0.11;
  Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 979 TGGGGATTGCTCTTACTTCTACGG 1004
  |||||||||||||||||||||||||||
Db 713 TGGGGATTGCTCTTACTTCTACGG 688
  |||||||||||||||||||||||||||

```

```

Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.
FEATURES
  source
    1..761
      /organism="Brassica oleracea"
      /mol_type="genomic DNA"
      /strain="TO1000DH3"
      /db_xref="taxon:3712"
      /clone="BONFA70"
      /clone_lib="BO_1.6_2_KB_tot"
      /note="Vector: pHOs1; Site 1: BstXI; 1.6-2 kb sheared
      total DNA inserted into pHOs1 using BstXI linkers"
ORIGIN
  Query Match 2.1%; Score 26; DB 28; Length 761;
  Best Local Similarity 100.0%; Pred. No. 0.11;
  Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 631 CAAGAAATCACTTCCCACTTGTGCT 656
  |||||||||||||||||||||||||||
Db 240 CAAGAAATCACTTCCCACTTGTGCT 265
  |||||||||||||||||||||||||||

RESULT 12
BH601155/c
LOCUS
DEFINITION
  BOHJF33TR BOHJ Brassica oleracea genomic clone BOHJF33, genomic
  survey sequence.
  824 bp DNA linear GSS 15-DEC-2001
ACCESSION
BH601155
VERSION
BH601155.1 GI:17853601
KEYWORDS
GSS.
SOURCE
  Brassica oleracea
  Brassica oleracea
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
  1 (bases 1 to 824)
REFERENCE
  Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
  Whole genome shotgun sequencing of Brassica oleracea
  Unpublished (2001)
  Other_GSSs: BOHJF33TF
  Contact: Chris Town
  TIGR
  9712 Medical Center Drive, Rockville, MD 20850, USA.
  Tel: 301-838-3523
  Fax: 301-838-0208
  Email: cdtown@tigr.org
  DNA is from a doubled haploid provided by Tom Osborn.
  Seq primer: TR
  Class: sheared ends.
FEATURES
  source
    1..824
      /organism="Brassica oleracea"
      /mol_type="genomic DNA"
      /strain="TO1000DH3"
      /db_xref="taxon:3712"
      /clone="BOHJF33"
      /clone_lib="BOHJ"
      /note="Vector: pHOs1; Site 1: BstXI; 2-3 kb sheared
      genomic DNA inserted into pHOs1 using BstXI linkers"
ORIGIN
  Query Match 2.1%; Score 26; DB 28; Length 824;
  Best Local Similarity 100.0%; Pred. No. 0.11;
  Matches 26; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
Qy 979 TGGGGATTGCTCTTACTTCTACGG 1004
  |||||||||||||||||||||||||||
Db 713 TGGGGATTGCTCTTACTTCTACGG 688
  |||||||||||||||||||||||||||

```

```

RESULT 13
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Homo sapiens (human)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
  1 (bases 1 to 737)
  Mahairas,G.G., Wallace,J.C., Smith,K., Swartzell,S., Holzman,T.,
  Keller,A., Shaker,R., Furlong,J., Young,J., Zhao,S., Adams,M.D. and
  Hood,L.
  Sequence-tagged connectors: A sequence approach to mapping and
  scanning the human genome
  Proc. Natl. Acad. Sci. U.S.A. 96 (17), 9739-9744 (1999)
  10449764
  Contact: Mahairas GG, Wallace JC, Hood L
  High Throughput Sequencing Center
  University of Washington
  401 Queen Anne Avenue North, Seattle, WA 98109, USA
  Tel: (206) 616-3618
  Fax: (206) 616-3887
  Email: jwallace@u.washington.edu
  Clones are derived from the human BAC library RPCI-11. For BAC
  library availability, please contact Pieter de Jong
  (pieter@dejong.med.buffalo.edu). Clones may be purchased from
  BACPAC Resources (http://bacpac.med.buffalo.edu/ordering_bac.htm)
  or from Resear h Genetics (info@resgen.com). BAC end Web Server:
  http://www.htsc.washington.edu
  Plate: 9002 row: C column: 21
  Seq primer: T7
  Class: BAC ends
  High quality sequence stop: 737.
  Location/Qualifiers
    1..737
      /organism="Homo sapiens"
      /mol_type="genomic DNA"
      /db_xref="taxon:9606"
      /clone="plate=9002 Col=21 Row=C"
      /sex="male"
      /clone_lib="RPCI-11 Human Male BAC Library"
      /notes="Vector: pBACE3.6; Site 1: EcoRI; Site 2: EcoRI;
      Male blood DNA was isolated from one randomly chosen donor
      and partially digested with a combination of EcoRI and
      EcoRI Methylase. Size selected DNA was cloned into the
      pBACE3.6 vector at EcoRI sites"
  ORGANISM
    Query Match
    Best Local Similarity 100.0%; Pred. No. 1.2;
    Matches 24; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

  QY 1164 TTGTGTTTTTTTGTATGTTCTT 1187
    |||||||
  Db 519 TTGTGTTTTTTTGTATGTTCTT 542
    |||||||

RESULT 14
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Arabidopsis thaliana (thale cress)
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
  1 (bases 1 to 445)
  Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R.,
  Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
  Shinn,P., Zimmermann,J. and Ecker,J.R.
  A Sequence-Indexed Library of Insertion Mutations in the
  Arabidopsis Genome
  Unpublished (2001)
  Contact: Joseph R. Ecker
  Salk Institute Genomic Analysis Laboratory (SIGAL)
  The Salk Institute for Biological Studies
  10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
  Tel: 858 453 4100 x1752
  Fax: 858 558 6379
  Email: ecker@salk.edu
  This is single pass sequence recovered from the left border of
  TDNA.
  Class: TDNA tagged.
  Location/Qualifiers
    1..445
      /organism="Arabidopsis thaliana"
      /mol_type="genomic DNA"
      /strain="Columbia 0"
      /db_xref="taxon:3702"
      /clone="SALK_011524"
      /clone_lib="Arabidopsis thaliana TDNA insertion lines"
      /notes="PCR was performed on Arabidopsis thaliana lines
      each of which contains one or more TDNA insertion
      elements. The resultant fragment for each line was
      directly sequenced to determine the genomic sequence at
      the site of insertion. Details of the protocols used can
      be found at http://signal.salk.edu/tdna_protocols.html"
  ORIGIN
    Query Match
    Best Local Similarity 100.0%; Pred. No. 3.3;
    Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1172 TTTTGTATTGTTCTTGTGATA 1194
      |||||||
    Db 54 TTTTGTATTGTTCTTGTGATA 76
      |||||||

RESULT 15
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Pan troglodytes (chimpanzee)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
  1
  Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
  Totoki,Y., Watanabe,H. and Sakaki,Y.
  BAC end sequences of Library PTB
  Unpublished
  2 (bases 1 to 1082)
  Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
  Totoki,Y., Watanabe,H. and Sakaki,Y.
  Direct Submission
  Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
  and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
  1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
  (E-mail:chimpes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
  Tel:81-45-503-9111, Fax:81-45-503-9170)
  Clones are derived from the chimpanzee BAC library PTB This BAC end
  was generated during the R&D process and may have higher chance of

```

```

ORGANISM
  Arabidopsis thaliana
  Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
  Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
  rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsi.
  1 (bases 1 to 445)
  Alonso,J.M., Leisner,T.J., Barajas,P., Chen,H., Cheuk,R.,
  Gadrinab,C., Jeske,A., Karnes,M., Kim,C.J., Parker,H., Prednis,L.,
  Shinn,P., Zimmermann,J. and Ecker,J.R.
  A Sequence-Indexed Library of Insertion Mutations in the
  Arabidopsis Genome
  Unpublished (2001)
  Contact: Joseph R. Ecker
  Salk Institute Genomic Analysis Laboratory (SIGAL)
  The Salk Institute for Biological Studies
  10010 N. Torrey Pines Road, La Jolla, CA 92037, USA
  Tel: 858 453 4100 x1752
  Fax: 858 558 6379
  Email: ecker@salk.edu
  This is single pass sequence recovered from the left border of
  TDNA.
  Class: TDNA tagged.
  Location/Qualifiers
    1..445
      /organism="Arabidopsis thaliana"
      /mol_type="genomic DNA"
      /strain="Columbia 0"
      /db_xref="taxon:3702"
      /clone="SALK_011524"
      /clone_lib="Arabidopsis thaliana TDNA insertion lines"
      /notes="PCR was performed on Arabidopsis thaliana lines
      each of which contains one or more TDNA insertion
      elements. The resultant fragment for each line was
      directly sequenced to determine the genomic sequence at
      the site of insertion. Details of the protocols used can
      be found at http://signal.salk.edu/tdna_protocols.html"
  ORIGIN
    Query Match
    Best Local Similarity 100.0%; Pred. No. 3.3;
    Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

    QY 1172 TTTTGTATTGTTCTTGTGATA 1194
      |||||||
    Db 54 TTTTGTATTGTTCTTGTGATA 76
      |||||||

RESULT 15
LOCUS
DEFINITION
ACCESSION
VERSION
KEYWORDS
SOURCE
ORGANISM
  Pan troglodytes (chimpanzee)
  Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
  Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Pan.
  1
  Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
  Totoki,Y., Watanabe,H. and Sakaki,Y.
  BAC end sequences of Library PTB
  Unpublished
  2 (bases 1 to 1082)
  Fujiyama,A., Hattori,M., Toyoda,A., Taylor,T.D., Yada,T.,
  Totoki,Y., Watanabe,H. and Sakaki,Y.
  Direct Submission
  Submitted (02-AUG-2001) Asao Fujiyama, The Institute of Physical
  and Chemical Research (RIKEN), Genomic Sciences Center (GSC);
  1-7-22 Suehiro-chou,Tsukumi-ku, Yokohama, Kanagawa 230-0045, Japan
  (E-mail:chimpes@gsc.riken.go.jp, URL:http://hgp.gsc.riken.go.jp/,
  Tel:81-45-503-9111, Fax:81-45-503-9170)
  Clones are derived from the chimpanzee BAC library PTB This BAC end
  was generated during the R&D process and may have higher chance of

```

Mon Sep 27 08:51:13 2004

clone tracking errors.

PRIMERS

Sequencing: -21M13

LIBRARY

Vector : pKS145

R.Site 1 : SacI

R.Site 2 : SacI.

Location/Qualifiers

1. .1082

/organism="pan troglodytes"

/mol\_type="genomic DNA"

/db\_xref="taxon:9598"

/clone="PTB-073J18.F"

/sex="male"

/cell\_type="lymphoblast"

/clone\_lib="PTB Chimpanzee Male BAC Library"

FEATURES

source

ORIGIN

Query Match 1.9%; Score 23; DB 29; Length 1082;

Best Local Similarity 100.0%; Pred.No. 4.4;

Matches 23; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 912 TTTCGCCGTCGTTGTTCCGGG 934

Db 678 TTTCGCCGTCGTTGTTCCGGG 656

Search completed: September 25, 2004, 14:17:27

Job time : 3727 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2004, 06:59:56 ; Search time 548 Seconds  
(without alignments)  
9496.428 Million cell updates/sec

Title: US-09-913-767-1

Perfect score: 1225

Sequence: 1 aaacgcgaagcagcaagaa.....aattgaaagatatgagct 1225

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 3373863 seqs, 2124099041 residues

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : N\_Geneseq\_29Jan04:\*  
1: Geneseqn1980s:\*  
2: Geneseqn1990s:\*  
3: Geneseqn2000s:\*  
4: Geneseqn2001as:\*  
5: Geneseqn2001bs:\*  
6: Geneseqn2002s:\*  
7: Geneseqn2003as:\*  
8: Geneseqn2003bs:\*  
9: Geneseqn2003cs:\*  
10: Geneseqn2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	1220.2	99.6	1225	3	AAA97919
2	901.2	73.6	2586	3	AAC44184
3	523.2	42.7	1049	3	AAA97920
4	450.6	36.8	493	3	AAC36831
5	143.6	11.7	2660	3	AAC47974
6	140.6	11.5	1071	3	AAA97925
7	135.8	11.1	1081	3	AAA97924
8	132.8	10.8	1418	3	AAC46395
9	131.2	10.7	1421	3	AAC39779
10	128.8	10.5	1403	3	AAC37652
11	127.2	10.4	1293	3	AAA97922
12	127.2	10.4	3387	6	AB214407
13	127.2	10.4	3387	7	ADAC6246
14	123.4	10.1	1200	7	ADA70627
15	117.2	9.6	83698	6	ABN85767
16	113.4	9.3	1194	3	AAA97923
17	103.8	8.5	1152	7	ADA70213
18	93.2	7.6	1145	3	AAA97921
19	93.2	7.6	1292	3	AAC51119
20	91.6	7.5	1295	3	AAC37289
21	86.8	7.1	2175	7	ADA70748
22	82	6.7	1155	7	ADA69686
23	74.2	6.1	1040	7	ADA69568

24	55	4.5	592	7	ABX56692	Arabidops
25	49	4.0	83391	6	ABQ67093	Abq67093 Human ang
26	48	3.9	14006	6	ABL33959	Ab133959 Human imm
27	47.2	3.9	17869	6	ABK39921	Abk39921 Human che
28	47.2	3.9	17869	6	ABL32105	Ab132105 Human imm
29	45.6	3.7	2000	6	ABZ15497	Abz15497 Arabidops
30	45.6	3.7	2000	7	ADA71938	Ada71938 Rice gene
31	45.6	3.7	7814	4	AAS46530	Aas46530 Tumour su
32	45.6	3.7	15674	6	ABL32362	Ab132362 Human imm
33	45.6	3.7	15674	6	ABL34476	Ab134476 Human met
34	45.6	3.7	15674	6	ABL70513	Ab170513 Chemocall
35	45.2	3.7	5192	6	ABL32232	Ab132232 Human imm
36	45.2	3.7	5312	6	ABK33960	Abk33960 Human DNA
37	45.2	3.7	5312	7	ADA20357	Ada20357 Prostate
38	45.2	3.7	5312	7	ADA84164	Ada84164 Human ren
39	45.2	3.7	5312	9	ADB54090	Adb54090 Pretreate
40	45.2	3.7	5312	9	ADB54218	Adb54218 Pretreate
41	45.2	3.7	5312	9	ADE84080	Ade84080 Human lym
42	45.2	3.7	5312	9	ADE84156	Ade84156 Human lym
43	45.2	3.7	14353	6	ABL34435	Ab134435 Human imm
44	45	3.7	4428	6	ABL34460	Ab134460 Human met
45	45	3.7	8294	6	ABL34430	Ab134430 Human imm

## ALIGNMENTS

## RESULT 1

AAA97919  
ID AAA97919 standard; DNA; 1225 BP.

XX AC AAA97919;

DT 19-JAN-2001 (first entry)

XX DE A. thaliana PUP1 DNA #1.

XX KW PUP1; transgenic plant; nucleobase transporter; apical dominance;

XX KM flowering behaviour; senescence; pesticide distribution; ds.

XX OS Arabidopsis thaliana.

XX FH Key Location/Qualifiers

FT CDS 25..1095

FT /\*tag= a

FT /product= "PUP1"

XX DE19907209-A1.

XX PD 24-AUG-2000.

XX PP 19-FEB-1999; 99DE-01007209.

XX PR 19-FEB-1999; 99DE-01007209.

XX PA (FROM/) FROMMER W.

XX PI Gillissen B, Buerkle L, Andre B, Frommer WB;

XX DR WPI; 2000-566202/53.

XX PS P-PSDB; AAB10684.

XX PT Nucleic acid, useful for producing transgenic plants with altered

XX PT nucleobase transport, encodes a nucleobase transporter protein of

XX PT Arabidopsis thaliana.

XX CC Claim 2; Page 11; 24pp; German.

XX CC This invention describes a novel nucleic acid encoding a plant nucleobase

XX CC transporter (I). (I) is produced by complementation of a nucleobase

XX CC transport (NBT)-defective host cell with a plant gene bank by selection

XX CC of NBT-positive cells. (I) is used to isolate homologous sequences from

XX CC bacteria, fungi, plants, animals and humans, for expression of the



PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 02-AUG-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 06-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 08-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 26-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 22-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.

[illegible]



us-09-913-767-1.rng

Mon Sep 27 08:51:14 2004

PR	24-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.	Query Match	36.8%;	Score 450.6;	DB 3;	Length 493;
PR	28-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.	Best Local Similarity	99.1%;	Pred. No. 1.9e-114;		
PR	29-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.	Matches 453;	Conservative	0;	Mismatches	4;
PR	30-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.					
PR	01-JUL-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.	QY	1	AAAAAGCAAGCAGCAAGAGATGAAGATGTTTGTATATCATATAAAGCTGTATTATC	60	
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.	DB	37	AAAACAGCAAGCAGCAAGAGATGAAGATGTTTGTATATCATATAAAGCTGTATTATC	96	
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.	QY	61	CTCACTATAGGACATGAGAGTCTTTGTTAACTCGTCTCTACTTTACCAATGGCGGA	120	
PR	08-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.	DB	97	CTCACTATAGGACATGAGAGTCTTTGTTAACTCGTCTCTACTTTACCAATGGCGGA	156	
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.	QY	121	AAACGAATCTGGTTCATGAGCTTCTATCAAACCGCTGGTTCATATCATCTCATCCCT	180	
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.	DB	157	AAACGAATCTGGTTCATGAGCTTCTATCAAACCGCTGGTTCATATCATCTCATCCCT	216	
PR	12-JUL-1999;	99US-0143477P.	PR	29-SEP-1999;	99US-0156596P.	QY	181	CTCTTGGTCTCTCTCTCAGCGGTCCGCGCAACCCCAACCCCTTAAACAACGCGGAAAC	240	
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.	DB	217	CTCTTGGTCTCTCTCTCAGCGGTCCGCGCAACCCCAACCCCTTAAACAACGCGGAAAC	276	
PR	14-JUL-1999;	99US-0143624P.	PR	06-OCT-1999;	99US-0157865P.	QY	241	AAGCGGAAACAAAGCTCTTCTCATGGAACCTCTCTGTGGAGCGCTCCATTGTGATA	300	
PR	15-JUL-1999;	99US-0144005P.	PR	07-OCT-1999;	99US-0158029P.	DB	277	AAGCGGAAACAAAGCTCTTCTCATGGAACCTCTCTGTGGAGCGCTCCATTGTGATA	336	
PR	16-JUL-1999;	99US-0144082P.	PR	08-OCT-1999;	99US-0158232P.	QY	301	GGTTCTCAGAGACTTGACAACTTATATTCTTACGGATTAGCATATCTGCCAGTT	360	
PR	16-JUL-1999;	99US-0144086P.	PR	12-OCT-1999;	99US-0158369P.	DB	337	GGTTCTCAGAGACTTGACAACTTATATTCTTACGGATTAGCATATCTGCCAGTT	396	
PR	19-JUL-1999;	99US-0144325P.	PR	13-OCT-1999;	99US-0159293P.					
PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159294P.					
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159295P.					
PR	19-JUL-1999;	99US-0144333P.	PR	14-OCT-1999;	99US-0159329P.					
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159330P.					
PR	20-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159331P.					
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.					
PR	20-JUL-1999;	99US-0144884P.	PR	18-OCT-1999;	99US-0159584P.					
PR	21-JUL-1999;	99US-0145086P.	PR	21-OCT-1999;	99US-0160741P.					
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.					
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.					
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160770P.					
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160814P.					
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160815P.					
PR	22-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160980P.					
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.					
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160989P.					
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161404P.					
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161405P.					
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161406P.					
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161359P.					
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161360P.					
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161361P.					
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161920P.					
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161922P.					
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.					
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.					
PR	04-AUG-1999;	99US-0147302P.								
PR	05-AUG-1999;	99US-0147192P.								
PR	05-AUG-1999;	99US-0147200P.								
PR	06-AUG-1999;	99US-0147303P.								
PR	06-AUG-1999;	99US-0147416P.								
PR	09-AUG-1999;	99US-0147935P.								
PR	10-AUG-1999;	99US-0148171P.								
PR	11-AUG-1999;	99US-0148319P.								
PR	12-AUG-1999;	99US-0148341P.								
PR	13-AUG-1999;	99US-0148565P.								
PR	13-AUG-1999;	99US-0148684P.								
PR	16-AUG-1999;	99US-0149368P.								
PR	17-AUG-1999;	99US-0149175P.								
PR	18-AUG-1999;	99US-0149426P.								
PR	20-AUG-1999;	99US-0149722P.								
PR	20-AUG-1999;	99US-0149723P.								
PR	20-AUG-1999;	99US-0149929P.								
PR	23-AUG-1999;	99US-0149902P.								
PR	23-AUG-1999;	99US-0149930P.								
PR	25-AUG-1999;	99US-0150566P.								
PR	26-AUG-1999;	99US-0150884P.								
PR	27-AUG-1999;	99US-0151065P.								
PR	27-AUG-1999;	99US-0151068P.								
PR	27-AUG-1999;	99US-0151080P.								
PR	30-AUG-1999;	99US-0151303P.								
PR	31-AUG-1999;	99US-0151438P.								
PR	01-SEP-1999;	99US-0151930P.								

QY 361 TCACTTCATCGCTCATATAATCGGAACCTCAACTAGCTTTCAACGCTCTCTTCGCTTTCTTG 420  
Db |||||  
QY 397 TCAACTTCATCGCTCATATAATCGGAACCTCAACTAGCTTTCAACGCTCTCTTCGCTTTCTTG 456  
Db |||||  
QY 421 TTAGTCAAGCAAAAGTTCACTCCGTTCTCCATAAAG 457  
Db 457 TTAGTCAAGCAAAAGTTCACTCCGTTCTCCATAAAG 493

RESULT 5  
AAC47974  
ID AAC47974 standard; DNA; 2660 BP.  
XX AAC47974;  
AC  
XX  
XX 18-OCT-2000 (first entry)  
XX  
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 55799.  
DE  
XX  
XX Hybridisation assay; genetic mapping; gene expression control;  
KW protein identification; signal transduction pathway; metabolic pathway;  
KW promoter; termination sequence; ss.  
XX  
XX Arabidopsis thaliana.  
XX  
XX EP1033405-A2.  
PN  
XX  
PD 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 28-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 02-JUL-1999; 99US-0142154P.  
PR 06-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 26-JUL-1999; 99US-0145224P.  
PR 27-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.





[illegible]

Mon Sep 27 08:51:14 2004

Db	250	TACGTAGTCTGGACCTCTTGTAGGACGAGATTGCTATCTGTACTCCATTGGACTCTTT	309	XX	06-SEP-2000.
Qy	349	TATCTGCCAGTTTCAACTTTCATCGTCATATCGGAATCAACTAGCTTTTCAACGCTCTC	408	XX	25-FEB-2000; 2000EP-00301439.
Db	310	TACTTACCCTGTTTACCTATCTCCCTGATCTGTGCACTCAGTTAGCTTCAATGCTTTC	369	XX	25-FEB-1999; 99US-0121825P.
Qy	409	TTCCGCTTCTTTGTAGTCAAGCAAAAGTTCACTCCGTTCTCCATAACGCCGCTGTTTG	468	PR	99US-0123180P.
Db	370	TTCTCTTATTTTCTTAACTCACAAAACCTTACCCTATCAATTTTAAATTTCTTTTCCCT	429	PR	99US-0133548P.
Qy	469	TTGACGGTTGGTATCGGATCCTTGGTTCACAGTGTAGGACAAACCGGCTAAGGAG	528	PR	99US-0125788P.
Db	430	CTAAGTATATCTTCCACCTACTTGCATCTCAATATGAGGACAGACTCCACAAAAGTT	489	PR	99US-0136264P.
Qy	529	ACCAAGAAAGATGTGTGGTTTGGGTTCTTGATGACTGTGGTTGAGCTTCTCTATGCT	588	PR	99US-0136785P.
Db	490	ACAAAGGAGATGTCAAGGTTTCATATGCCCGTTGCTGCTGTGTTATGGT	549	PR	99US-0127462P.
Qy	589	TTTATATACCGTCTGTGAGTCACTTACAAGAACTCGTCAAGAAATCACTTTCCCA	648	PR	99US-0128234P.
Db	550	CTAGTCTTATCCCTACACAGTAGCCCTTTCTAAAGTCTTAAAGACCAAAATTTCTCA	609	PR	99US-0128714P.
Qy	649	CTTGTCTTGGAGATTGAGATGCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	708	PR	99US-0129845P.
Db	610	GAAGTTATGATATGATATCTACGTGAGTCTAGTGGCCAGTTGTGTAGCGTGGTGGG	669	PR	99US-0130077P.
Qy	709	ATGTTTCATCGTGGAGATTTTAAGGTGATAGCAAGAGCAAGAGTTCAAGATTGGA	768	PR	99US-0130449P.
Db	670	CTTTTGTCTAGCAGTGGTGAAGTCTTTGAGCAGTGA--AATGATTAATCAAAAT	726	PR	99US-0130510P.
Qy	769	GGATCAGTGTTTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATG	828	PR	99US-0130891P.
Db	727	GGGAAGTATCTTACATATTAACCTAGTGTGACAGCTGTTACCTGGCAGGTATTTCTCC	786	PR	99US-0131449P.
Qy	829	TTAGGACCTAGGATGTTGTTTGTGATCATCATCTAGCTCTGCTGCTGCTGCTGCTGCTG	888	PR	99US-0132048P.
Db	787	ATCGTGGCAGCAGTGTCTTCGAGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT	846	PR	99US-0132485P.
Qy	889	GTTCTGCTTCGGTGACTGAAGTTTTCGCGTCTGTTGTTTTCGGGAGAGTTTTCAGGCA	948	PR	99US-0132863P.
Db	847	TTGGGACTCCAGGTTTCTTATCTTGGCGTATCATTTTCCATGACAAATGAATGGG	906	PR	99US-0134256P.
Qy	949	GAGAAAGTGTCTCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1008	PR	99US-0134218P.
Db	907	TTAAGTGTATTTCTATGATCTAGTCTTATTTGGGTTTCTACTTCTTCTTCTTCTTCTTCT	966	PR	99US-0134219P.
Qy	1009	TTTAAATCCGCAAGAAAGTTGTTGATTAACCTCAACCGCGGAGACAGAACTGCTTATT	1068	PR	99US-0134370P.
Db	967	TATCTTGTATGACAAAACCTTGAAGAAAATCATGAAAATCAACACAGAAATCCCTGAC	1026	PR	99US-0134768P.
Qy	1069	CTTCCAG 1075		PR	99US-0134941P.
Db	1027	CCACAG 1033		PR	99US-0135124P.
RESULT 8				PR	99US-0135353P.
AAC46395				PR	99US-0136021P.
ID AAC46395 standard; DNA; 1418 BP.				PR	99US-0136392P.
XX				PR	99US-0136782P.
XX AAC46395;				PR	99US-0137222P.
XX				PR	99US-0137528P.
18-OCT-2000 (first entry)				PR	99US-0137724P.
Arabidopsis thaliana DNA fragment SEQ ID NO: 49985.				PR	99US-0138094P.
Hybridization assay; genetic mapping; gene expression control;				PR	99US-0138540P.
KW protein identification; signal transduction pathway; metabolic pathway;				PR	99US-0138847P.
KW promoter; termination sequence; ss.				PR	99US-0139119P.
XX Arabidopsis thaliana.				PR	99US-0139452P.
XX				PR	99US-0139453P.
EP1033405-A2.				PR	99US-0139454P.
				PR	99US-0139455P.
				PR	99US-0139456P.
				PR	99US-0139457P.
				PR	99US-0139458P.
				PR	99US-0139459P.
				PR	99US-0139460P.
				PR	99US-0139461P.
				PR	99US-0139462P.
				PR	99US-0139463P.
				PR	99US-0139750P.
				PR	99US-0139763P.
				PR	99US-0139817P.
				PR	99US-0139899P.
				PR	99US-0140353P.
				PR	99US-0140354P.
				PR	99US-0140695P.
				PR	99US-0140823P.
				PR	99US-0140991P.
				PR	99US-0141287P.
				PR	99US-0141842P.
				PR	99US-0142154P.

PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143824P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144085P.	PR	07-OCT-1999;	99US-0158029P.
PR	19-JUL-1999;	99US-0144325P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144331P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144814P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0145086P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160741P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160767P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145089P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145192P.	PR	21-OCT-1999;	99US-0160814P.
PR	23-JUL-1999;	99US-0145145P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-0145224P.	PR	22-OCT-1999;	99US-0160981P.
PR	26-JUL-1999;	99US-0145276P.	PR	22-OCT-1999;	99US-0160989P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145951P.	PR	25-OCT-1999;	99US-0161406P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161359P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-0146388P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146389P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149902P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
<hr/>					
QY	283	AGCGCTCCATTGTCATAGGGTGTCTCAGAGGACTTGCACAACTTGTGACAACTTATATATCTTACGGA	342	22-SEP-1999;	99US-0155139P.
DB	429	ATCTTCTCGTGTGTGATCGGGCTTGTCTCGGTTTCAACAACTTCTTCTCATGGGA	488	23-SEP-1999;	99US-0155486P.
QY	343	TTAGCATATCTGCCAGTTTCAACTTTCATCGCTCATATATCGGAACTCAACTAGCTTTCAAC	402	24-SEP-1999;	99US-0155659P.
DB	489	ACCTCGTACTTCCGGTGTCCAGTCTATCGTCTCTCTCGACACAACTCGTTCCTCACT	548	28-SEP-1999;	99US-0156458P.
QY	403	GCTCTCTTCTGCTTCTTGTGTAGTCAAGAAAGTTCACTCGGTTCTCCATTAACGGCGTC	462	29-SEP-1999;	99US-0156596P.
DB	549	CTCAATTTGTCTAGGATCATAGTGAACAGAGATCACTTCTTCAAACTTCAACTCGGTT	608	04-OCT-1999;	99US-0157117P.
QY	463	GTCTTCTGTCAGCGTGTGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT	522	05-OCT-1999;	99US-0157753P.
DB	609	GTCTTCTGTCAGCGTGTGATCGGATCGGATCGGATCGGATCGGATCGGATCGGATCGGAT	668	06-OCT-1999;	99US-0157865P.
QY	523	AAGGAGCAAGAAAGATGATGTTGGTGTCTTGTATGATGATGATGATGATGATGATGATGAT	582	07-OCT-1999;	99US-0158029P.
DB	669	GGTTTAAACCAACCAAGTATTTTCATCGGGTATGATTCACGATCGGAGCGGTTTACTC	728	08-OCT-1999;	99US-0158232P.
QY	583	TATGCTTTTATATTAACCGTCTGTTAGCTTAACCTTAACGAAAGCTCGTCAAGAAATCACT	642	12-OCT-1999;	99US-0158369P.
DB	729	TTGCTCTTTTACCTCC-----CGTACGCGAGAGCTATACCGTACCGTTTATTTGT	779	13-OCT-1999;	99US-0159293P.
QY	643	TTCCCACTTGTGCTTTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAGATTCAG	702	13-OCT-1999;	99US-0159294P.
DB	780	TACGCGATGCTCATGAGGTCACCTGGTGTGAAATTCGACGACGGTTCGCGACA	839	13-OCT-1999;	99US-0159295P.
QY	703	ATTGGCATGTTTCATCGTTCGAGATTTTAAAGGTATAGCAAGAGAGAGAGTTTCAAG	762	14-OCT-1999;	99US-0159329P.

Mon Sep 27 08:51:14 2004

Db	840	ATCGTATGGCTTGGGAAGCGGTTTAAGGAAATGGTTAAGGAAGCAACCATGTTTC	18-MAY-1999;	99US-0134768P.
Qy	763	ATTGGAGGATCAGTGTTTTACTATGCATT---GATAGTGATCACAGGAATAATATGGCAA	19-MAY-1999;	99US-0134941P.
Db	900	ACCAAAGGACCAACATTTTATTTGACCGTTCGCGATACCTAGCAATGTTGGTGCAGTGGCAG	20-MAY-1999;	99US-0135124P.
Qy	820	GGTTTCTTCTTAGGAGCCATAGGAGATTGTGTTTGTGTCATCATCCTAGCTTCTGTGTTT	21-MAY-1999;	99US-0135353P.
Db	960	CTCTCTTTTCGAGCCACGCTCAGGATGTTTACTTGACGTGAGTATCACCGAGGTATC	24-MAY-1999;	99US-0135629P.
Qy	880	CTGATAAGTGTCTCTTCGGTGACTGAAGTTTTCGCCGTCGTTTTCCTCCGGGAGAAG	25-MAY-1999;	99US-0136021P.
Db	1020	TGCATGACGGCGTTCCTCCGATGAATCTGATAGAGGTGGTGGCTTACGGTGACGTG	27-MAY-1999;	99US-0136392P.
Qy	940	TTTCAGGACAGAAAGGTGCTCTACTTCTTCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	28-MAY-1999;	99US-0136782P.
Db	1080	TTTGTGGAGTGAAGATTGTGTCGACGGTGCTATGTATTTGGGATCTCATCTTACACA	01-JUN-1999;	99US-0137222P.
Qy	1000	TACGGCGAGTTTAAATCCGGCAAGAAAG	03-JUN-1999;	99US-0137528P.
Db	1140	TACGGGATGTATGAGATGAAGAAGG	04-JUN-1999;	99US-0137528P.
RESULT 9				
AAC39779				
ID	AAC39779 standard; DNA; 1421 BP.			
AC	AAC39779;			
XX	17-OCT-2000 (first entry)			
DE	Arabidopsis thaliana DNA fragment SEQ ID NO: 25972.			
XX	Hybridisation assay; genetic mapping; gene expression control;			
KW	protein identification; signal transduction pathway; metabolic pathway;			
KW	promoter; termination sequence; ss.			
XX	Arabidopsis thaliana.			
XX	EP1033405-A2.			
XX	06-SEP-2000.			
XX	25-FEB-2000; 2000EP-00301439.			
XX	25-FEB-1999; 99US-0121825P.			
PR	05-MAR-1999; 99US-0123180P.			
PR	09-MAR-1999; 99US-0123548P.			
PR	23-MAR-1999; 99US-0125788P.			
PR	25-MAR-1999; 99US-0126264P.			
PR	29-MAR-1999; 99US-0126785P.			
PR	01-APR-1999; 99US-0127462P.			
PR	06-APR-1999; 99US-0128234P.			
PR	08-APR-1999; 99US-0128714P.			
PR	16-APR-1999; 99US-0129845P.			
PR	19-APR-1999; 99US-0130077P.			
PR	21-APR-1999; 99US-0130449P.			
PR	23-APR-1999; 99US-0130510P.			
PR	28-APR-1999; 99US-0130891P.			
PR	30-APR-1999; 99US-0131449P.			
PR	30-APR-1999; 99US-0132048P.			
PR	04-MAY-1999; 99US-0132407P.			
PR	05-MAY-1999; 99US-0132484P.			
PR	06-MAY-1999; 99US-0132485P.			
PR	06-MAY-1999; 99US-0132486P.			
PR	07-MAY-1999; 99US-0132487P.			
PR	11-MAY-1999; 99US-0132863P.			
PR	14-MAY-1999; 99US-0134218P.			
PR	14-MAY-1999; 99US-0134219P.			
PR	14-MAY-1999; 99US-0134221P.			
PR	14-MAY-1999; 99US-0134370P.			

PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145931P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 23-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 23-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 23-OCT-1999; 99US-0161404P.  
PR 23-OCT-1999; 99US-0161405P.  
PR 23-OCT-1999; 99US-0161406P.

PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.  
  
Query Match 10.7%; Score 131.2; DB 3; Length 1421;  
Best Local Similarity 50.5%; Pred. No. 1.1e-25;  
Matches 378; Conservative 0; Mismatches 358; Indels 12; Gaps 2;  
  
QY 283 AGCGCTCCATTGTCATAGGGTTGCTCACAGACTTGCTCACTCACTCGGTTCTCCATAAAGCGGTC 462  
DB 429 ATCTTCTCCGTTGATCGGCTTGTTCGGTTTCAACAACCTTCTCTCATGGGA 488  
QY 343 TTAGCATATCTGCCAGTTTCAACTTCATCGCTCATATCGGAATCGGAATCACTTCTCAACCTCACTGCTTCAAC 402  
DB 489 ACCTGTTACCTTCCGGTGTCCAGTTCATCGCTTCTCTCTCGACAACTCGTCTTCACT 548  
QY 403 GCTCTCTTCCGTTTCTTGTAGTCAAGCAAAAGTTCACTCGGTTCTCCATAAAGCGGTC 462  
DB 549 CTCATTTTGTCTAGGATCATTTGTAACAGAAATCACTTCTCAACCTCACTGCTTCACT 608  
QY 463 GTTTTGTGACGGTTGATTCGGGATCTTTCGCTTACAGTGTATGGAGACAACCGGCT 522  
DB 609 GTCTCTGTGACGTTAAGCTCTGTTTGTGGCTTCTGATTCGAGTAAAGATAAACCCTCC 668  
QY 523 AAGGAGACAAGAAAGATATGTTGGTGTCTTCTGATGACTGCTGTTGCTGAGCTTCTC 582  
DB 669 GGTAAACCAAAACCAAGTATTTTCATCGGTTATGATTCACGATCGGAGCGGTTTACTC 728  
QY 583 TATGCTTTTATATACCGCTCGTTGAGCTAACTTACAAGAAAGCTCGTCAAGAAATCACT 642  
DB 729 TTGCTCTTTTACCTCCC-----CGTACCGAGAGCTATACCGTACCGTTTATGT 779  
QY 643 TTCCCACTTGTGCTTGAGATTACAGATGGTCAATGCTTGTCTGCTACTTTTTTCTGTGTC 702  
DB 780 TACCGCATGCTCATGAGGTGCAACTGGTGTGATGGAATTCGACGACGGTTTTTCGGGACA 839  
QY 703 ATTGGCATGTTTCATCGTTGGAGATTAAAGTGTATGACGAGAAAGCAAGAGAGATTCAAG 762  
DB 840 ATCGGTATGGCTTTCGAAAGCGGGTTTAAAGGAAATGTTAAGGAAGCAACCATGTTTC 899  
QY 763 ATTGGAGGATCAGTGTTTTACTATGCACT---GATAGTGTATCACAGAAATATGCA 819  
DB 900 ACCAAAGGACCAACATTTTATGGACGTTTGCATACTAGCAATGTGGTACGTCGAG 959  
QY 820 GGTTCCTTTAGGAGCCATAGGATGTTGTTTGTGCATCATCACTAGCTTCTGTTGTT 879  
DB 960 CTCTCTTTCGACCCACGTCAGGATGTTTACTTGCACGTCAGGTATCACCGGAGGTATC 1019  
QY 880 CTGATAGTGTCTGCTTCCGGTGAAGTTTTCGCGTCTGTTTGTTCGGGAGAG 939  
DB 1020 TGCATGACGCGGTGCTCGCATGAATGTGATAGGAGGTGTGGTGTCTTACGTTGACGTG 1079  
QY 940 TTTTCAGGAGAGAAAGGTGCTCTCTACTTCTTCTTCTTGGGGATTTGCTCTTACTTC 999  
DB 1080 TTTGGTGGAGTGAGATTGTGTCGACGGTGTATGTTTGGGGATTTCTCATCTTACACA 1139  
QY 1000 TACGCGAGTTTAAATCCGGCAAGAAAG 1027  
DB 1140 TACGGGATGTATGAAGATGAAGAAG 1167  
  
RESULT 10  
AAC37652  
ID AAC37652 standard; DNA; 1403 BP.  
XX  
AC AAC37652;  
XX  
DT 17-OCT-2000 (first entry)  
XX

us-09-913-767-1.rng

Mon Sep 27 08:51:14 2004

```
DE Arabidopsis thaliana DNA fragment SEQ ID NO: 18172.
XX Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX
OS Arabidopsis thaliana.
XX EP1033405-A2.
XX
PD 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 06-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 07-MAY-1999; 99US-0132487P.
XX 11-MAY-1999; 99US-0132863P.
XX 14-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137502P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140655P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 20-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 21-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 22-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 27-JUL-1999; 99US-0145951P.
PR 28-JUL-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 02-AUG-1999; 99US-0147038P.
PR 03-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 09-AUG-1999; 99US-0148171P.
PR 10-AUG-1999; 99US-0148319P.
PR 11-AUG-1999; 99US-0148341P.
PR 12-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
```





1159 TCAATTGTCTATCAGCACTACCTCGACGAAAGATTGAAGACTAGCCACACAAGTCCT 1218

1051 GAGACAGAACTGCGCTATTCTTCAGATTAGTGA 1082  
1219 GTAGGAGATCCTCATCTACTACTGCTGAGGA 1250

RESULT 12  
ABZ14407  
ID ABZ14407 standard; DNA; 3387 BP.

AC ABZ14407;  
XX 21-JAN-2003 (first entry)

XX Arabidopsis thaliana stress regulated gene SEQ ID NO 2212.  
XX Arabidopsis thaliana; plant; gene; stress; transgenic; ds.

OS Arabidopsis thaliana.  
XX  
XX W0200216555-A2.  
XX 28-FEB-2002.

XX 24-AUG-2001; 2001WO-US026685.  
XX 24-AUG-2000; 2000US-0227866P.  
XX 26-JUN-2001; 2001US-0264647P.  
XX 22-JUN-2001; 2001US-0300111P.

XX (SCRI ) SCRIPPS RES INST.  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX Harper JF, Kreps J, Wang X, Zhu T;  
XX WPI; 2002-304127/34.

DR Identifying a stress condition to which a plant cell has been exposed and  
XX producing plants with increased tolerance to these abiotic stresses.  
XX Claim 144; SEQ ID NO 2212; 577pp + Sequence Listing; English.

XX The invention relates to identifying a stress condition to which a plant  
XX cell has been exposed, comprising: (a) contacting nucleic acid  
XX representative of expressed polynucleotides in the plant cell with an  
XX array or probes representative of the plant cell genome; and (b)  
XX detecting a profile of expressed polynucleotides in the plant cell  
XX characteristic of a stress response. The method is useful in the  
XX production of transgenic plants, cells and seeds and in producing plants  
XX with increased tolerance to abiotic stress. The present sequence is that  
XX of an Arabidopsis thaliana stress regulated gene (ABZ12196-ABZ17574) used  
XX in methods of the invention. Note: The sequence data for this patent is  
XX not represented in the printed specification but is based on sequence  
XX information supplied to Derwent by the European Patent Office

XX Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;  
SQ

Query Match 10.4%; Score 127.2; DB 6; Length 3387;  
Best Local Similarity 48.2%; Pred. NO. 2e-24; Indels 3; Gaps 1;  
Matches 391; Conservative 0; Mismatches 418

QY 271 ACTCTCTGTGAGCGCTCCATTCATAGGGTTGCTCAGAGACTTGACAATCTTA 330  
DB 376 ACCATTCTTGGATCAGTTTACATCGTTACTGACTATTAGTGTCTGTAATCTTATATG 435  
QY 331 TATTTCTTACGATTAGCATATCTGCGAGTTTCAGTTTCATCGTCATAATCGGAATCAA 390  
DB 436 TCCTCTGTGGTTTACTATACTACTACTGAGTTTCTATCTTCTCCCTCATCTTGGCTCAAA 495  
QY 391 CTAGCTTTCAACGCTCTCTTCGCTTTTCTTGTAGTCAAGCAAAAGTTCACTCGGTTCTCC 450  
DB 496 TTGGCCCTTCACTGCGCTTCTTCTCATATTTCCTAAATTCGAGAAAGTTTCAACCTTTTCA 555

XX Claim 1f; Page 13; 24pp; German.

XX This invention describes a novel nucleic acid encoding a plant nucleobase  
XX transporter (I). (I) is produced by complementation of a nucleobase  
XX transport (NBT)-defective host cell with a plant gene bank by selection  
XX of NBT-positive cells. (I) is used to isolate homologous sequences from  
XX bacteria, fungi, plants, animals and humans, for expression of the  
XX encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting  
XX expression of (II) (when in antisense orientation), and to produce  
XX transgenic crop plants. The transgenic plants have modified nucleobase  
XX transport properties, e.g. altered affinity and substrate specificity  
XX that may result in more efficient nucleobase transport in leaves, changes  
XX in apical dominance, flowering behaviour and senescence, or improved  
XX distribution of pesticides. This sequence encodes the Arabidopsis  
XX thaliana PUP1 protein which is described in the method of the invention

XX Sequence 1293 BP; 345 A; 288 C; 247 G; 413 T; 0 U; 0 Other;  
SQ

Query Match 10.4%; Score 127.2; DB 3; Length 1293;  
Best Local Similarity 48.2%; Pred. No. 1.3e-24;  
Matches 391; Conservative 0; Mismatches 418; Indels 3; Gaps 1;

QY 271 ACTCTCTGTGAGCGCTCCATTCATAGGGTTGCTCAGAGACTTGACAATCTTA 330  
DB 442 ACCATTCTTGGATCAGTTTACATCGTTTACTGACTATTAGTGTCTGTAATCTTATATG 501

QY 331 TATTTCTTACGATTAGCATATCTGCGAGTTTCAGTTTCATCGTCATAATCGGAATCAA 390  
DB 502 TCCTCTGTGGTTTACTATACTACTACTGAGTTTCTATCTTCTCCCTCATCTGSCCTCAAA 561

QY 391 CTAGCTTTCAACGCTCTCTTCGCTTTTCTTGTAGTCAAGCAAAAGTTCACTCGGTTCTCC 450  
DB 562 TTGGCTTCACTGCTCTTCTCATATTCTTAAACTCGCAGAAGTTTCAACCTTTCAAT 621

QY 451 ATAAAGCCGCTGTTTGTGACGGTTGGTATCGGATCCTTCGGTTTACACAGTATGGA 510  
DB 622 GTGAATTTCTGTTTCTCTTACTATTCTCTGCGCTTCTCTGTTCAACTGATTCG 681

QY 511 GACAAACCGCTAAGGAGACAAAGAGTATGTGGTTGGGTTCTTGTGATGACTGTGGTT 570  
DB 682 GAAACACAGCAAAAGTGTCTAGATTAATATGTATAGGGATAATATGTACCATTTGT 741

QY 571 GCAGCTCTTCTATGCTTTTATATACCGCTGTTGAGTACTTACAGAAAGCTCGT 630  
DB 742 GCTTCTGCTGGATGGATGCTGCTATCCCTGCTACACTGATCTCAGGAAGTTTGA 801

QY 631 CAAGAAATCACTTCCCACTGCTTGAGATTGAGATGCTATGCTGCTTGTCTACT 690  
DB 802 AAGAAAGCAAACTTCTCAACGGTCACTGCTTGGTCCCTTACCAATCTCTAGTTGCAAG 861

QY 691 TTTTCTGCTGTCATGGCAGTTTCACTCGTTGGAGATTTAAGGTGATGACAGAGAGCA 750  
DB 862 TGTGTGTTCTCATAGGACTTTTCGCAAGCGGGAGTGGAAAATTTTAACTAGATG 921

QY 751 AGAGAGTTCAAGATGGAGATCAGTGTCTTACTATGATTTGATGATGATGATGATGAT 810  
DB 922 GAAAACTCAAACT---GGGGAAGATGCGCATGCTTATGACTTTGGCTCGATGCTATT 978

QY 811 ATATGGCAAGGTTTCTTCTTAGGAGCCATAGGGAATGTTTGTGATCATCATCTAGCT 870  
DB 979 TCTTGGCAAGTCTACACCAATGGCGTCTGGGAGTGTATCTTGAAGTCTCTCTGTTTC 1038

QY 871 TCTGTTGTTCTGATTAAGTTTCTGCTCGGTGAGTGAAGTTTTCGCGCTGTTTGTTC 930  
DB 1039 TCCAAATCCATAAATCTGCTGCGGATGTTGCTATAGTTCCAGTTGTAGCAGTATTTTC 1098

QY 931 CGGAGAGATTTTCAGGACAGAAAGGTTGTCTCTTACTTCTTCTTCTTTTGGGATTTGTC 990  
DB 1099 CATGATAAATGACCGCTCAAGATCTTCTCCATCAATTTAGCTATCTGCGGATTCATT 1158

QY 991 TCTTACTTCTACGGGAGTTTAAATCCCGCAGAAAGTTGTTGATAAACTCAACCGCG 1050

QY 451 ATAAACCCGCTCGTTTGGTGGATCGGATCCTTGGTTACAGTGATGGA 510  
Db |||||  
QY 556 GTGAATCTCTGTTTCTCTTACTATTTCTCTGCCCTCTCGGGTCAACACTGATTCG 615  
Db |||||  
QY 511 GACAAACCGCTAGGAGACAGAAAGATGATGTTGGTCTTGTGACATGTTGGTT 570  
Db |||||  
QY 616 GAAACACACAGCAAAAGTCTAGAGTAAATATGTATAGGGATAATATGTACATTTGGT 675  
QY 571 GCAGCTCTCTCTATGCTTTTATATACCGCTCGTTGAGCTAACTTACAAGAAAGCTCGT 630  
Db |||||  
QY 676 GCTTCTGCTGGGATGGATTGCTGCTATCTCTGTACACTGATCTCTCAGAAAGTTTAA 735  
QY 631 CAAGAATACATTTCCACTTGTCTGATGATTCAGATGTCATGTCCTTGTCTGCTACT 690  
Db |||||  
QY 736 AAGAAGCAAACTTCTCAACCGTCACTGACTTGTCTGCTTACCAATCTCTAGTTGCAAGC 795  
QY 691 TTTTCTGCTGCTATGCTGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 750  
Db |||||  
QY 796 TGTGTTCTCTATAGACTTTTTCGACGCGGAGTGGAAACTTTTAAACAAGTGAGATG 855  
QY 751 AGAGAGTTCAAGATTGAGGATCAGTGTCTTATCTATGATGATGATGATGATGATGAT 810  
Db |||||  
QY 856 GAAACTTACAACT---GGGAAAGTCCATACGTTATGACTTTGGCTCGATGATTT 912  
QY 811 ATATGGCAAGTTCTTTTAGGACCAATAGGATGTTTGTGTCATGATGATGATGATGAT 870  
Db |||||  
QY 913 TCCTGGCAAGTCTACCACTTGGCTGCTGGGACTGATCTTTGAGTCACTCTCTGTTGTT 972  
QY 871 TCTGGTGTCTGATAGTGTCTGCTTCCGCTGACTGAACTTTTCCGCTGCTGTTTCTTC 930  
Db |||||  
QY 973 TCCAAATCCAACTGCTGTGGGATGCTTATAGTTCCAGTTGTAGCAGTGTGTTTCTTC 1032  
QY 931 CGGAGAAAGTTTTCAGGACAGAAAGTGTCTCTACTTCTTCTTCTTCTTCTTCTTCTTCT 990  
Db |||||  
QY 1033 CATGATAAATGAACGGTCAAGATCTTCTCCATCATTTTGTAGTATCTGGGATTCATT 1092  
QY 991 TCTTACTCTPACGGCGAGTTTAAATCGGCAAGAAAGTTGTTGATAAACCTCAACCGCG 1050  
Db |||||  
QY 1093 TCATTGTCTATCAGCACTACCTCGACGAAAGAAAGTTGAAGACTAGCCACACAGTCT 1152  
QY 1051 GAGACAGACTGCTTCTTCTCAGTTAGTGA 1082  
Db |||||  
QY 1153 GTAGGAGATCCTCATCTACTACTCTGCTGAGGA 1184

## RESULT 13

ADA68246  
ID ADA68246 standard; DNA; 3387 BP.  
XX  
AC ADA68246;  
XX  
XX  
DT 20-NOV-2003 (first entry)  
XX  
DE Arabidopsis thaliana gene, SEQ ID 908.  
XX  
XX Plant; bacterial infection; fungal infection; viral infection; ds.  
XX  
XX Arabidopsis thaliana.  
OS  
XX  
XX WO2003000898-A1.  
PN  
XX  
XX 03-JAN-2003.  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
PF  
XX  
XX 22-JUN-2001; 2001WO-IB001105.  
PR  
XX  
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.  
XX  
XX Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;  
PI Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;  
PI  
XX

DR  
XX  
PT  
PT  
PT  
PT  
XX  
PS  
XX  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
CC  
SQ

WPI; 2003-175290/17.

Identifying at least one gene involved in plant resistance or response to pathogenic infection for conferring resistance or tolerance to a plant to bacterial, fungal or viral infection by determining or detecting plant gene expression.

Claim 6; SEQ ID NO 908; 899pp; English.

The present invention relates to a method (M1) for identifying genes involved in plant resistance or response to pathogenic infection. M1 comprises identifying a gene whose expression is significantly altered in the incompatible interaction of plant gene expression relative to expression of the gene in an uninfected plant, in a mutant plant that does not express a gene associated with response to pathogenic infection, or in a corresponding incompatible or compatible interaction. (M1) is useful for conferring resistance to resistance or tolerance to a plant to bacterial, fungal or viral infection. The present sequence was used to illustrate the invention.

Sequence 3387 BP; 870 A; 781 C; 654 G; 1082 T; 0 U; 0 Other;

Query Match 10.4%; Score 127.2; DB 7; Length 3387;  
Best Local Similarity 48.2%; Pred. No. 2e-24;  
Matches 391; Conservative 0; Mismatches 418; Indels 3; Gaps 1;

QY 271 ACTCTCTGTGGAGCGCTCCATTGTGATAGGGTGTCTCACAGGACTTGCACAACTACTTA 330  
Db |||||  
QY 376 ACCATTCTTGGATCAGTTTACATCGTTACTGGACTATTAGTGTCTGCTAACTCTTATG 435  
QY 331 TATTCTTACGATAGCATATCTGCCAGTTTCAACTTCATCGCTCATATAATCGGAACCTCA 390  
Db |||||  
QY 436 TCCTCTGTTGGTTTACTATCTTACCAGTTTCTTCTCTCTCTCTCTCTCTCTCTCTCT 495  
QY 391 CTAGCTTTCAACGCTCTCTCTGCTTTCTTGTAGTCAAGCAAAAGTTCACTCGCTTCTCC 450  
Db |||||  
QY 496 TTGGCTTCTCACTGCTCTCTCTCTATATTTCTAACTCGCAGAAAGTTTCAACCTTTCA 555  
QY 451 ATAAACCGCTGCTTGTGAGGTTGTATCGGGATCCTTCCGTTTACACAGTGTATGGA 510  
Db |||||  
QY 556 GTGAATCTCTGTTCTCTCTTCTTCTTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 615  
QY 511 GACAAACCGCTAAGGAGACAGAAAGATATGTGTTGGTTCCTTGTGATGATGATGATG 570  
Db |||||  
QY 616 GAAACACACAGCAAAAGTGTCTAGGTAAATATGTATAGGGATAATATGTACCATTTGGT 675  
QY 571 GCAGCTCTCTCTATGCTTTTATATACCGCTGTTGAGCTAACTTACAAGAAAGCTCGT 630  
Db |||||  
QY 676 GCTTCTGCTGGGATTTGGATTGCTGCTATCTCTGTTACAACTGATCTCTCAGGAAGTTT 735  
QY 631 CAAGAAATCACTTTCCCACTTGTGCTTGTAGATTTCAGATGCTATGCTGCTGCTACT 690  
Db |||||  
QY 736 AAGAAGCAAACTTCTCAACGGTCACTGACTTGGTCTGTACCAATCTCTAGTTGCAAGC 795  
QY 691 TTTTCTGTGTCAATGGCATGTTTCTATGCTTGTGAGATTTTAAAGGTGATAGCAAGAGA 750  
Db |||||  
QY 796 TGTGTTGTTCTCATAGACTTTTTCGCAAGCGGAGTGGAAACTTTAACAAGTGTAGATG 855  
QY 751 AGAGATTCAAGATTGGAGATCAGTGTGTTTCTATGATGATGATGATGATGATGATGAT 810  
Db |||||  
QY 856 GAAACTTACAACT---GGGAAAGTGCATACGTTATGACTTTGGCTCGATGATTT 912  
QY 811 ATATGGCAAGTTTCTTCTTAGGACCAATAGGATGTTTGTGTCATGATGATGATGATG 870  
Db |||||  
QY 913 TCCTGGCAAGTCTACCACTTGGCTGCTGGGACTGATCTTTGAGTCACTCTCTGTTGTT 972  
QY 871 TCTGGTGTCTGATAGTGTCTGCTTCCGCTGACTGAACTTTTCCGCTGCTGTTTCTTC 930  
Db |||||  
QY 973 TCCAAATCCAACTGCTGTGGGATGCTTATAGTTCCAGTTGTAGCAGTGTGTTTCTTC 1032  
QY 931 CGGAGAAAGTTTTCAGGACAGAAAGTGTCTCTACTTCTTCTTCTTCTTCTTCTTCTTCT 990  
Db |||||  
QY 1033 CATGATAAATGAACGGTCAAGATCTTCTCCATCATTTTGTAGTATCTGGGATTCATT 1092  
QY 991 TCTTACTCTPACGGCGAGTTTAAATCGGCAAGAAAGTTGTTGATAAACCTCAACCGCG 1050  
Db |||||  
QY 1093 TCATTGTCTATCAGCACTACCTCGACGAAAGAAAGTTGAAGACTAGCCACACAGTCT 1152  
QY 1051 GAGACAGACTGCTTCTTCTCAGTTAGTGA 1082  
Db |||||  
QY 1153 GTAGGAGATCCTCATCTACTACTCTGCTGAGGA 1184

QY	991	TCCTACTTCTACGGCGAGTTTAAATCCGGCAGAAAGTTGTTGATAAACCTCAACCGCG	1050
Db	1093	TCATTTGTTCTATCAGCACTACCTCGACGAAAGAAAGTTGAAGACTACCCACAAAGTCTCT	1152
QY	1051	GAGACAGAACTGCTTATCTTCCAGTTAGTGA	1082
Db	1153	GTAGGAGATCTCTATCTACTACTGCTGAGGA	1184
RESULT 14			
AD70627	ADA70627 standard; DNA; 1200 BP.		
XX	AC	ADA70627;	
XX	20-NOV-2003	(first entry)	
XX	DE	Rice gene, SEQ ID 3950.	
XX	KW	Plant; bacterial infection; fungal infection; viral infection; rice;	
XX	KW	gene; ds.	
XX	OS	Oryza sativa.	
XX	XX	WO200300898-A1.	
XX	PN	03-JAN-2003.	
XX	PD	22-JUN-2001; 2001WO-IB001105.	
XX	PF	22-JUN-2001; 2001WO-IB001105.	
XX	PR	(SYGN ) SYNGENTA PARTICIPATIONS AG.	
XX	PA	Chang H, Chen W, Cooper B, Glazebrook J, Goff SA, Hou Y;	
XX	PI	Katagiri F, Quan S, Tao Y, Whitham S, Xie Z, Zhu T, Zou G;	
XX	PI	WPI; 2003-175290/17.	
XX	DR	Identifying at least one gene involved in plant resistance or response to	
XX	XX	pathogenic infection for conferring resistance or tolerance to a plant to	
XX	PT	bacterial, fungal or viral infection by determining or detecting plant	
XX	PT	gene expression.	
XX	XX	Claim 6; SEQ ID NO 3950; 899pp; English.	
XX	PS	The present invention relates to a method (M1) for identifying genes	
XX	CC	involved in plant resistance or response to pathogenic infection. M1	
XX	CC	comprises identifying a gene whose expression is significantly altered in	
XX	CC	the incompatible interaction of plant gene expression relative to	
XX	CC	expression of the gene in an uninfected plant, in a mutant plant that	
XX	CC	does not express a gene associated with response to pathogenic infection,	
XX	CC	or in a corresponding incompatible or compatible interaction. (M1) is	
XX	CC	useful for conferring resistance to resistance or tolerance to a plant to	
XX	CC	bacterial, fungal or viral infection. The present sequence was used to	
XX	CC	illustrate the invention.	
XX	SQ	Sequence 1200 BP; 143 A; 416 C; 434 G; 206 T; 0 U; 1 Other;	
Query Match			
Best Local Similarity 10.1%; Score 123.4; DB 7; Length 1200;			
Matches 408; Conservative 49.1%; Pred. No. 1.4e-23;			
Matches 408; Indels 27; Gaps 2;			
QY	258	CTTCTCATGGAACCTCTCTGTGAGCGGCTCCATTCTCATAGGTTGCTCAGAGCT	317
Db	330	CCTCTCTCTCCGCGCGCTCTGTCGCGCGCGCGCGCTCTGCGCGGCTGTACGCGGT	389
QY	318	TGACAACTACTTATATCTTACGAGATTAGCATATCTGCCAGTTTCAACTTCATCGCTCAT	377
Db	390	GTGTCGCTTCGTGTACGCGCTGGGTCCGAGCGCTGCGCGCTGTCCAGTGTGCTGCT	449
QY	378	AATCGGAACCACTAGCTTTCAACGCTCTCTTCGCTTTCTTGTAGTCAAGAAAGTT	437

Db	450	GCTGCGCAGCAGCTGGCTTTCACGCGGTTTCGCGTCTCTCTCTGCGGCTCCGCTT	509
QY	438	CACCTCGGTTCTCCATAAACCGCGTCTGTTTGTGACCGGTTGATCGGATCCTTTCGTT	497
Db	510	CACGCGGTTCTCGGCAACCGCGTCTGCTCTCACMATCGGCGCGGCTGCTGGCGGT	569
QY	498	ACACAGTGTGAGACAAACCGGCTAAGGAGAGCAAGAAAGATATGTGTGTTGGTTCTT	557
Db	570	CGGCGCGTCTGCGGGAAGCGCGGCGGATCTCTCCAGGCGTACTCGACGGGTTCTG	629
QY	558	GATGACTGTGTTGAGCTCTTCTCTATGCTTTTATATATTACCGTCTGTTAGCTTA	617
Db	630	CGAGGCGATCGCGCGCGGCGGTAGCGGCTGTGTGATCCCGCTCTGTCGAGGTGCGCAC	689
QY	618	CAAGA-----AAGCTGTCAAGAAATCACTTCCCACTTGT	653
Db	690	GCGAGGTACGGGCGCGCACCGGCGCGCGGAGGTGCGCGCTCTCTACGCGACGCT	749
QY	654	GCTTGAGATTACAGTGTCTATGCTTGTCTGTACTTTTCTGTGTCTATTTGGCATGTT	713
Db	750	GATGCGATGCGCGGCGGTGATGCGCGCGGCGCACGCGGTGTCGCTGCTCGCATGCG	809
QY	714	CATCGTTGGAGATTTAAGTGTATGACAGAGAGAGAGTTCAAGATTGGAGGATC	773
Db	810	GATCAAGGCGCACTTCAGGCGGTGCGCGGGAAGCGCGGCTTCGGGCTCGCGCGGC	869
QY	774	AGTGTTTTACTGCTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG	833
Db	870	CAACTACTACTCTGCTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	926
QY	834	AGCATAGGAGTTGTTTGTGTCATCATCATCATCATCATCATCATCATCATCATCATCAT	893
Db	927	CATCATGGGCTCATCATCATCATCATCATCATCATCATCATCATCATCATCATCATCAT	986
QY	894	GTTTCCGCTGACTGAAGTTTTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG	953
Db	987	CCTGCCCTCTCGCAGGTCTCTCGCGCTCTCGCGCTCTCGCGCTCTCGCGCTCTCGCG	1046
QY	954	AGTGTCTCTCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT	1013
Db	1047	GGGATCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT	1106
QY	1014	ATCCGCAAGAAAGTTGTTGATAACTCAACCGCGGAGACAGAACTGCC	1064
Db	1107	GCAGAGAAGAAAGAGGCGGAGAGATGCGCGAGCGCGAGAGGTGCG	1157
RESULT 15			
ABN85767	ID ABN85767 standard; cDNA; 83698 BP.		
XX	AC	ABN85767;	
XX	21-OCT-2002	(first entry)	
XX	DE	Arabidopsis yellow stripe1-like 4 encoding cDNA SEQ ID NO 9.	
XX	KW	Maize; transgenic; plant; yellow stripe1-like; ysl; Arabidopsis;	
XX	KW	iron uptake; bioremediation; yellow stripe 1; ysl; gene; ss.	
XX	OS	Arabidopsis sp.	
XX	PN	WO200240698-A2.	
XX	PD	23-MAY-2002.	
XX	PF	16-NOV-2001; 2001WO-US043101.	
XX	PR	16-NOV-2000; 2000US-0249222P.	
XX	PA	(UYA ) UNIV YALE.	
XX	XX		



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:47:26 ; Search time 23.3672 Seconds  
(without alignments)  
777.685 Million cell updates/sec

Title: US-09-913-767-9  
Perfect score: 1760  
Sequence: 1 MVALVINCIIAIGNCGS.....QQESQTEQSSLRPISEC 352

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA:\*  
1: /cgn2\_6/ptodata/2/iaa/5A COMB.pdp.\*  
2: /cgn2\_6/ptodata/2/iaa/5B COMB.pdp.\*  
3: /cgn2\_6/ptodata/2/iaa/6A COMB.pdp.\*  
4: /cgn2\_6/ptodata/2/iaa/6B COMB.pdp.\*  
5: /cgn2\_6/ptodata/2/iaa/PGTUS COMB.pdp.\*  
6: /cgn2\_6/ptodata/2/iaa/backfiles1.pdp.\*

Pred. No. is the number of results predicted by chance to have a  
score greater than or equal to the score of the result being printed,  
and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	129	7.3	327	4	US-09-352-991A-25058
2	126	7.2	292	2	US-09-024-848-2
3	126	7.2	292	3	US-09-348-116A-2
4	126	7.2	494	4	US-09-543-681A-5990
5	125.5	7.1	464	4	US-09-252-991A-25514
6	118.5	6.7	514	4	US-09-543-681A-4255
7	118.5	6.7	867	4	US-09-107-532A-4393
8	114.5	6.5	296	4	US-09-540-236-2922
9	114	6.5	325	4	US-09-489-039A-8024
10	114	6.5	405	4	US-09-543-681A-4311
11	112.5	6.4	469	4	US-09-328-352-5007
12	112.5	6.4	730	4	US-09-328-352-4442
13	112	6.4	434	4	US-09-328-352-7304
14	112	6.4	503	4	US-09-134-001C-3948
15	110.5	6.3	462	2	US-08-998-376-2
16	110.5	6.3	462	2	US-08-998-376-4
17	110.5	6.3	940	4	US-09-328-352-8165
18	110	6.2	414	4	US-09-489-039A-12578
19	109.5	6.2	424	4	US-09-543-681A-7510
20	109	6.2	297	4	US-09-134-001C-5605
21	108.5	6.2	303	4	US-09-252-991A-27266
22	108.5	6.2	445	4	US-09-328-352-28655
23	107.5	6.1	449	4	US-09-328-352-7512
24	107.5	6.1	720	4	US-09-328-352-19581
25	107	6.1	394	4	US-09-134-000C-3732
26	107	6.1	421	4	US-09-543-681A-7791
27	107	6.1	465	4	US-09-198-452A-524

28	106.5	6.1	444	4	US-09-328-352-5249	Sequence 5249, Ap
29	106.5	6.1	447	4	US-09-252-991A-28081	Sequence 28081, A
30	106.5	6.1	494	4	US-09-134-001C-4475	Sequence 4475, Ap
31	106	6.0	308	4	US-09-328-352-7092	Sequence 7092, Ap
32	105.5	6.0	451	4	US-09-328-352-7659	Sequence 7659, Ap
33	105.5	6.0	516	4	US-09-543-681A-5385	Sequence 5385, Ap
34	105.5	6.0	934	4	US-09-252-991A-19515	Sequence 19515, A
35	103.5	5.9	233	4	US-09-024-848-4	Sequence 4, Appli
36	103.5	5.9	233	3	US-09-348-116A-4	Sequence 4421, Ap
37	103.5	5.9	616	4	US-09-543-681A-4421	Sequence 17668, A
38	103	5.9	317	4	US-09-252-991A-17668	Sequence 5653, Ap
39	103	5.9	977	4	US-09-134-000C-5653	Sequence 3001, Ap
40	102.5	5.8	496	4	US-09-134-001C-3001	Sequence 7419, Ap
41	102.5	5.8	552	4	US-09-489-039A-7419	Sequence 4163, Ap
42	102	5.8	696	4	US-09-107-532A-4163	Sequence 8204, Ap
43	101.5	5.8	463	4	US-09-543-681A-8204	Sequence 24, Appli
44	101.5	5.8	503	3	US-09-068-195-24	Sequence 4689, Ap
45	101.5	5.8	536	4	US-09-328-352-4689	

ALIGNMENTS

RESULT 1  
US-09-252-991A-25058  
; Sequence 25058, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; FILE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 25058  
; LENGTH: 327  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-25058

Query Match	7.3%	Score 129;	DB 4;	Length 327;
Best Local Similarity	24.1%	Pred. NO. 4.8e-05;		
Matches	69;	Conservative	47;	Mismatches 114;
				Indels 56;
				Gaps 11;
QY	41	LETAGFPVIFPILFSGYITRRSNNVGDSTSFLLKPRLLIAAVIVGILSGFD--NYLYA	98	
DB	65	LTAAGLRFLIAFPFLFLLFARLR-----HETIFF--PRLLAFVFTVLTGYFSVPYLLN	116	
QY	99	YGIAYLPVSTAAIIASQALAFIAIFSFVMVKKHFTPTTNAVLLTVGAAL-----G	151	
DB	117	YGRQYVASGLTALLFSSMPVILIFSAIFLRER---IYLSQILGIAIGFSGFLMIRSQG	173	
QY	152	MHTETDKPVHETHKQVITGFLITVAANVAFILPLVELAYQAKQTMSTVILEFQLIL	211	
DB	174	LHLDYTE-----LLGVLAAILAAIMHA-----LSYVITKRGADIGVITNTLP	217	
QY	212	CLLASIVSVIGMFIAQDFKQALPKAREFKGLGALFWAVFSAIIVQGFGLAIGLIFS	271	
DB	218	IGTAG---LGLPMAGLVMEHPNLSAVTSRWAAIIVLGLVAS-----VGPIVYFF	265	
QY	272	TSSLVSGIMISVLLPITEVLAVIF--YHEKFAEKLGLSLALSLMGF	315	
DB	266	LLKRLSPVLSFVFIIFPFAVIGAWYE-----GVAISRELIGF	305	

RESULT 2  
US-09-024-848-2  
; Sequence 2, Application US/09024848.

Mon Sep 27 08:51:16 2004

```

; Patent No. 5962295
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: No. 5962295el LicB
; NUMBER OF SEQUENCES: 4
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/024,848
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/033,807
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GMS0017
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-024-848-2

```

```

Query Match 7.2%; Score 126; DB 2; Length 292;
Best Local Similarity 23.2%; Pred. No. 8.4e-05;
Matches 67; Conservative 55; Mismatches 99; Indels 68; Gaps 15;

Qy 37 PSTLETAGFPV-----IFPLFSYITRRSRNNVGDSTFFFLKPRLLIAVIVGI 88
Db 28 PSITDLSPPVVAATHDFLSIFILLAFLLVKEKVR-----LSIFL---NIRNVSVIIGA 79

Qy 89 L-----SGPDNLYAYGIAIYLPVSTAALIIASQAFIAIFSFVVKHKFTPTTINAVLLT 144
Db 80 LLAGPIGQANLYA--VKYIGSSLASVSALYPAISVLLAFFFLKHKISKNTVFGIVLI- 136

Qy 145 VGAAVLGMHTTDPVHETHKQYITGELITVAARV-----MYAFILPLVELAY 192
Db 137 ----IGGIIATYK-VEQVNSFYI-GILCALVCAIANGSESVLSSPFAMESELSEIALLI 190

Qy 193 QKAKQMTSYTLVLEPQ-----LILCLLASIVSVIGMFIAGDFKQALPKEA 237
Db 191 RQVTSFSLVYLVLFSHQSFVAVANGQLGLMIVFVAFDMISYLAYIIA--INRLQPAKA 248

Qy 238 REFKLGEALFYVAVFSIAIWQGFPLG-AIGLIFSTSLV--SGIMISV 283
Db 249 ----TGLNVSVYV---WTVLFAALFLGTSLDMLTMTSLVVIAGVYIII 290

```

```

RESULT 3
US-09-348-116A-2
; Sequence 2, Application US/09348116A
; Patent No. 6200779
; GENERAL INFORMATION:
; APPLICANT: Lonetto, Michael A.
; TITLE OF INVENTION: No. 6200779el LicB
; NUMBER OF SEQUENCES: 4

```

```

; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Dechert, Price & Rhoads
; STREET: 4000 Bell Atlantic Tower, 1717 Arch Stre
; CITY: Philadelphia
; STATE: PA
; COUNTRY: USA
; ZIP: 19103-2793
; COMPUTER READABLE FORM:
; MEDIUM TYPE: Diskette
; COMPUTER: IBM Compatible
; OPERATING SYSTEM: DOS
; SOFTWARE: FastSEQ for Windows Version 2.0
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/348,116A
; FILING DATE:
; CLASSIFICATION:
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: US/09/024,848
; FILING DATE:
; APPLICATION NUMBER: 60/033,807
; FILING DATE: 28-FEB-1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Falk, Stephen T
; REGISTRATION NUMBER: 36,795
; REFERENCE/DOCKET NUMBER: GMS0017
; TELEPHONE: 215-994-2488
; TELEFAX: 215-994-2222
; TELEX:
; INFORMATION FOR SEQ ID NO: 2:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 292 amino acids
; TYPE: amino acid
; STRANDEDNESS: single
; TOPOLOGY: linear
;
US-09-348-116A-2

```

```

Query Match 7.2%; Score 126; DB 3; Length 292;
Best Local Similarity 23.2%; Pred. No. 8.4e-05;
Matches 67; Conservative 55; Mismatches 99; Indels 68; Gaps 15;

Qy 37 PSTLETAGFPV-----IFPLFSYITRRSRNNVGDSTFFFLKPRLLIAVIVGI 88
Db 28 PSITDLSPPVVAATHDFLSIFILLAFLLVKEKVR-----LSIFL---NIRNVSVIIGA 79

Qy 89 L-----SGPDNLYAYGIAIYLPVSTAALIIASQAFIAIFSFVVKHKFTPTTINAVLLT 144
Db 80 LLAGPIGQANLYA--VKYIGSSLASVSALYPAISVLLAFFFLKHKISKNTVFGIVLI- 136

Qy 145 VGAAVLGMHTTDPVHETHKQYITGELITVAARV-----MYAFILPLVELAY 192
Db 137 ----IGGIIATYK-VEQVNSFYI-GILCALVCAIANGSESVLSSPFAMESELSEIALLI 190

Qy 193 QKAKQMTSYTLVLEPQ-----LILCLLASIVSVIGMFIAGDFKQALPKEA 237
Db 191 RQVTSFSLVYLVLFSHQSFVAVANGQLGLMIVFVAFDMISYLAYIIA--INRLQPAKA 248

Qy 238 REFKLGEALFYVAVFSIAIWQGFPLG-AIGLIFSTSLV--SGIMISV 283
Db 249 ----TGLNVSVYV---WTVLFAALFLGTSLDMLTMTSLVVIAGVYIII 290

```

```

RESULT 4
US-09-543-681A-5990
; Sequence 5990, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05

```



```

: ||| ||| ||| : : : ||| : : |||
Db 193 IGKTLVGLVLPATAIAGLFGALVSRYPGTPSAELV-----EQIAHE 238

: ||| ||| ||| ||| : : : ||| : : |||
Qy 163 THRQYITGFLITVAAAVVMYAFILPL-----VELAVQKA 195

: ||| ||| ||| ||| ||| : ||| : |||
Db 239 PETQDLPSFGVTLATVLLPVFLMLLKTFFADVAFPDGHAFRAWMDMIGHPI SALLALLVA 298

: ||| ||| ||| ||| ||| ||| : ||| : |||
Qy 196 KQWMSYTLVLEFQILICLL-----ASIVSVIGMFIAGDPKQALPKEARFKLGE---- 244

: ||| ||| ||| ||| ||| ||| ||| : ||| : |||
Db 299 LYTFGYARGPELEKILRLDQSLAPAAI VMIIG--AGGGFKQMLVASGVGDVIGHLAVN 356

: ||| ||| ||| ||| ||| ||| ||| ||| : ||| : |||
Qy 245 -----ALFYVVAVFSAILIQGFLGAI---GLIFSTSSLVSGIMISVLLPITEVLAVIF 295

: ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| : |||
Db 357 AQISPILLAWLVAAVIRIATGSATVATIGAGIVVPVIDLPGVNNRELLVLATGAGSLIL 416

: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| : |||
Qy 296 YH-----EKFOAEKGLSLALS LSGFV 316

: ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| : ||| : |||
Db 417 SHVNDAGFWLVKQYFNMSVSETPKTWTAMETILSVVGLV 455

RESULT 6
US-09-543-681A-4255
; Sequence 4255, Application US/09543681A
; Patent No. 5605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEIN
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709.1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 4255
; LENGTH: 514
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-4255

```

Db	408	YVLLSDFSTFPVFLSILLILLIILAKSNELFREQLIYSWEWRTIDGIIIGALSLLY----	I 463
QY	276	VSG-----IMISVLLPTEV--LAVIVFH-----	EK 299
Db	464	VIGVNLPDFPHRRHHFTISFLPFSEKIWFSGLLAIIVSMVLVFLVHFLQEKKGIGEA	523
QY	300	FOAEKGLSLALSGCFVSFYFGEIKSG-----EKKRIQQEESOE	339
Db	524	FNEEKALKI-LITYG-----GNDSQLIIFLKDKMFAYEKDGE	560

## RESULT 8

```

US-09-540-236-2922
; Sequence 2922, Application US/09540236
; Patent No. 6673910
; GENERAL INFORMATION:
; APPLICANT: Gary L. Breton et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO
; TITLE OF INVENTION: FOR DIAGNOSTICS AND T
; FILE REFERENCE: 2709.2005-001
; CURRENT APPLICATION NUMBER: US/09/540,236
; CURRENT FILING DATE: 2000-04-04
; NUMBER OF SEQ ID NOS: 3840
; SEQ ID NO 2922
; LENGTH: 296
; TYPE: PRT
; ORGANISM: M.catarrhalis
US-09-540-236-2922

```

[illegible]

## RESULT 9

```

US-09-489-039A-8024
; Sequence 8024, Application US/09489039A
; Patent No. 6610836
; GENERAL INFORMATION:
; APPLICANT: Gary Breton et. al
; TITLE OF INVENTION: NUCLEIC ACID AND
; TITLE OF INVENTION: NUCLEIC ACID FOR DI
; FILE REFERENCE: 2709.2004001
; CURRENT APPLICATION NUMBER: US/09/489,
; CURRENT FILING DATE: 2000-01-27
; PRIOR APPLICATION NUMBER: US 60/117,7
; PRIOR FILING DATE: 1999-01-29
; NUMBER OF SEQ ID NOS: 14342
; SEQ ID NO 8024
; LENGTH: 325
; TYPE: prt
; ORGANISM: Klebsiella pneumoniae
US-09-489-039A-8024

```

```

RESULT 7
US-09-107-532A-4393
; Sequence 4393, Application US/09107532A
; Patent No. 6583275
; GENERAL INFORMATION:
; APPLICANT: Lynn A Doucette-Stamm and David Bush
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS
;
; NUMBER OF SEQUENCES: 7310
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION
; STREET: 100 Beaver Street
; CITY: Waltham
; STATE: Massachusetts
; COUNTRY: USA
; ZIP: 02154
;
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD-ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
;
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
;
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
;
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Denke
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
;
; INFORMATION FOR SEQ ID NO: 4393:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 867 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHEICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
;
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...867
; SEQUENCE DESCRIPTION: SEQ ID NO: 4393:
US-09-107-532A-4393

Query Match      6.7%; Score 118.5; DB 4; Length 867;
Best Local Similarity 19.8%; Pred. No. 0.0021;
Matches 80; Conservative 71; Mismatches 118; Indels 135; Gaps 22;

QY      49  IFIPILFYSYITRRSRNNVGDS-----SF-----FLIKPR-----78
DB      179  LYFPVVFLETFKKKGFLGTSLSKTQGLLVLSLEWTGVLFSLSVGYLMEIRIDLWQT 238
QY      79  --LLIAAVIGILSGPDNYLYAYGIAVLPVSTAALIASQLAFTAIPTSPFMVKKHPTPT 136
DB      239  IPLFAASVIGISMIPEIGSGPDMMI--IGLSAIGVPREIVVWIL--LYRLFYIIPFL 296
QY      137  INAVULL--TVGAAVLGMHETDKPV--HETHK-----QYITGFLITVAAMVAFILPLV 188
DB      297  IGIVFFPKNGSTPDQYSGIPKQLATEIAHKIVVVLLYFSGIMLVISAIPQAF-----T 352
QY      189  ELAVQAKQMTSYTLVLEFQLICLLASIVSGVMFIAGDFKQA--LPKEAREFKLGEALF 247
DB      353  EFRWLHSLNPLKHFHFIQFPSI--LLGFLIIMGRTAARVKRAYLP---TIFLIVLALL 407
QY      248  YVW-----AVFSAIL-----WQ-----GFFLGALGLIFSTSSL 275
DB      111  :::::

```

Query Match 6.5%; Score 114; DB 4; Length 325;  
Best Local Similarity 22.0%; Pred. No. 0.0016;  
Matches 80; Conservative 49; Mismatches 131; Indels 104; Gaps 16;

QY 2 VKALVIINCILIA-----IGNCGGLIMRLYFNNGGKRIWFSTFLETAGFPVFIPLIFS 56  
DB 26 IVALLIITILWAFSPSLIGE-----YLAGSVDSYFSLMRVGLAALVFLPPL-- 73  
QY 57 YITRRSNVNDSTSPFLIKPRLLIAAVIVGILSGFDNYLYAYGIAYLPVSTAALIASQ 116  
DB 74 --RTRQS-----PRIVLYMLVGMQ-----LGIWLLAFRAYLYLT-- 109  
QY 117 LAFIAPSPFWKHKTPPTINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVA 176  
DB 110 ---VSEFLFTV---FTPLYITLIYDLISG-----RKLRLWGYLLSAA 145  
QY 177 AAVMYAFILVLAYOKQATMSYTLVLEFQILCLLASIV-----SVIG 222  
DB 146 LAVLGRAI-----IRYDKVSDHFWTGLLLVQLSNLCEFAIGMVGKRLMEVRPMPQHNAFA 200  
QY 223 MFIAGDFKQALPKAREFKLGEA--LFYVAVFSAIIWQGFILGAIGLIE-----STSS 274  
DB 201 WFYVGAFIVAV---AAWFLGNAGKQLPTTLQHSVLVWLGVAASGLGYFWMNYGATQVDA 257  
QY 275 LVSGIMISVLPTTEVLAVIFYHEKFOAEKGLSLAL-----SLM---GFVSIFYGEIKSGE 327  
DB 258 GTLGINNVHVPAGLLVNLAIWQEQHPWPSFISGALVILASLVWHRRWVAPHSGQ--TAD 315  
QY 328 DKRR 331  
DB 316 DRRR 319

RESULT 10  
US-09-543-681A-4311  
; Sequence 4311, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; PRIOR FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 4311  
; LENGTH: 405  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-4311

Query Match 6.5%; Score 114; DB 4; Length 405;  
Best Local Similarity 20.5%; Pred. No. 0.0022;  
Matches 80; Conservative 54; Mismatches 136; Indels 120; Gaps 16;

QY 29 NNGGR-----IWFSTFLETAGPPVI--FIPL-----LFSY 57  
DB 10 NNSWKRNLIVYVWFGCELTGAGFSLIMPFLPYVEELGIKDHESLNLMTVGVAFTFLFSA 69  
QY 58 ITRRRSNVND--STSFLLKP-----RLIAAVIGILSGF---DN 94  
DB 70 IAPFWGKLSDRGRKMLRLSALGWAIVMVLGPAQNIWQLLILRALGLVGGFVFNAN 129  
QY 95 YLYA-----YGIAYLPVSTAAL-----IIASQLAFIAIF 123  
DB 130 ALIATQVPVKKSGWALGTSTGAVSGALIGLIGILLADLYLRPVFFITAAVLFCFIV 169  
QY 124 SFWVXHKFTPTINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVAANVYAF 183  
DB 190 TLPFVSENFTPVSKD-----ALSTQVFFSLKRLVICLFTTMIQVATGS,238

QY 184 ILPLVELAYOKQATMSYTLVLEFQILCLLASIVSGMFIAGDF-----KQALPKARE 239  
DB 239 VTILTYIYRDLAGSIS-----NLAFISGVIAVPGIAALISAPRFGLGDRIGPDKVLI 293  
QY 240 FKLGEALPYVAVFSAILI---WQ-----GFLGAIGLIFSTSSLSVGMISVLLPITEVLA 292  
DB 294 FTGLGSIFMLIPM--ALVSNYWEIGALRFLLGAV-----NAAMLPAVQTLILYINITPAIA 346  
QY 293 -VFYHEKFOAEKGLSLALSLSLWGFVSFYG 321  
DB 347 GRIFSYNQALRDVGNVTGMLMGAFVANYG 376

RESULT 11  
US-09-328-352-5007  
; Sequence 5007, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: GARY L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 5007  
; LENGTH: 469  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-5007

Query Match 6.4%; Score 112.5; DB 4; Length 469;  
Best Local Similarity 22.3%; Pred. No. 0.0038;  
Matches 74; Conservative 51; Mismatches 108; Indels 99; Gaps 19;

QY 44 AGPFIPIPLFSYITRRSNVNDSTSPFLIKPRLLIAAV-IVG-----ILSGFDNYL 96  
DB 157 AGSPLLFVVIYTSILR---GVGDSTT-----PLLALALTSVIGLVITPILLKY--- 203  
QY 97 YAGIAYLPVSTAALIIASQLAFIAIFSPFWKHKFTPTINAVLLTVGAALVGMHTET 156  
DB 204 --FGLPALGIIAPAIATITGYVAILFLAIYLNKXHLRPNRQLL----- 247  
QY 157 DKPVHETHKQYI-----TGF-LITVAAAVMYAFILPLVE-----LAYQKAKOT 198  
DB 248 ---QHHRNSELSEKIILRLGVPTGIONITTSWAGL--VIIGLVNHYGSHATAAYGAVNQV 302  
QY 199 MSYTLVLEFQILCLLASIVSGMFIAGDFKQALPKAREFKLG-EALF----- 247  
DB 303 LNY---IQPPAISISIAA--SIFAAQAIAGKSDLLARVTRTALGNWFLTGALIALGYL 357  
QY 248 ---YVAVFSA-----IIWQGFILGALIFSTSSLSVSG-IMISVLLPITE 289  
DB 358 FSKYLMALFITDPTVVVLGQQLLFIVLWAILFFGA-GAIFASIMRASGTVTVPMLINISA 416  
QY 290 VLAV-----IFYHEKFOAEKGLSLALSLSLWGFVS 317  
DB 417 ILFIEIPCAVWFSSMWGLKGIWVAYAL-AFVS 447

RESULT 12  
US-09-328-352-4442  
; Sequence 4442, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: GARY L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252

151 -GMHTETDKPVHETHKQYITGFLI--TVAAAVMYAFILPLVELAYOKAKQTMSTYLVLEF 207  
 176 APINHYNVPSE--GFVNGILTMDILGALV--FGIVIIQAIYSRG--VTDNKLVTYK 227  
 208 OLILCLLASIVSVMFIAGDFKQALPKAREFKLGEALFVVV--AVFSALIWQGF-- 262  
 228 AII-----ASLISGVGLTLV-----YLSLFLKGLGSHEVAFNAANGAVILHAYVOHA 274  
 263 LGAIGLIFSTSLVSGIMISVL 284  
 275 FGNMGSLEFLTGLIFLACMVTAI 296

RESULT 14  
 US-09-134-001C-3948  
 ; Sequence 3948, Application US/09134001C  
 ; Patent No. 6380370  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO STAPHYLOCOCCUS  
 ; TITLE OF INVENTION: EPIDERMIDIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC-007  
 ; CURRENT APPLICATION NUMBER: US/09/134,001C  
 ; PRIOR FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/064,964  
 ; PRIOR FILING DATE: 1997-11-08  
 ; PRIOR APPLICATION NUMBER: US 60/055,779  
 ; PRIOR FILING DATE: 1997-08-14  
 ; NUMBER OF SEQ ID NOS: 5674  
 ; SEQ ID NO 3948  
 ; LENGTH: 503  
 ; TYPE: PRT  
 ; ORGANISM: Staphylococcus epidermidis  
 US-09-134-001C-3948

Query Match 6.4%; Score 112; DB 4; Length 503;  
 Best Local Similarity 23.5%; Pred. No. 0.0047;  
 Matches 87; Conservative 53; Mismatches 105; Indels 126; Gaps 21;

71 SFPLIKPRLLIAAVIVGILSGFDNYLA-----YGIAY---LPVSTAALLIASOLA 118  
 164 SFILAWLSLLIAVILLVIATG-DWYIYSNNDDNAINYGMRLCINLLIHLAVIIPAAQFP 222  
 119 FIAIPFPMVKHETPTTINAVV---LLTVGAALVGHMTETDKPVHETHKQYITGFLITV 175  
 223 ----FOGMLIESVAAPTFSAIMHAGIVNAGVILLTRFS-----PVF-NDEIAISLLIIA 273  
 176 AAAYMYAFILPLVELAY--OKAKQTMSTYTLVLEFQLILCLL-----ASIVSI--GMFIAG 227  
 274 SISVLSSGSLTLVHVDYKQLVRSTISQ---MGFMVQCALGAYSAIVHLHLHGVEKAT 330  
 228 DFKQ-----ALPKEAREFKLGEALFYVVA-----VFSAILI--- 257  
 331 LFLQSGSVVKRFNIPTSPSVKSYGMLVFGRLAILLAIIFWLNDRHVDVLSALILAW 390  
 258 -----WQ---GFFLGAIGLI-----FSTSS-----LVSGIMISVL 284  
 391 SLUMSVNQIVAFSHGLIGRIVGCMIIIVAVIITHHYPTTILSNVDIHIIVSPPLISII 450  
 285 LPITEVLAVIFYHEKFOAEKGLSIALSLW-----GFVSIFYGEIKSGDKERRIQEE 336  
 451 LSI-----AIIVF-----GSMLSIIVWSRRRESKAFKLYLWLKVGKAK-----TQ 491  
 337 SOETQSSLSR 347  
 492 SIESHPSYLKR 502

RESULT 15  
 US-08-998-976-2  
 ; Sequence 2, Application US/08898976  
 ; Patent No. 5891670  
 ; GENERAL INFORMATION:

12 VALVIINCIIAIGNCGPLIMRLYFNNGG-----KRIWFSTFETAGFPVIFIPILFS 56  
 39 IPILLIFSLSLWLGEGADIKSAVTFEFSWAALGYSFKEFWAPLIDE-----LPVPVLTK 92  
 57 YITRRSNVNGSTSFELIKPRLIIAAVIGILSGFDNYLYAYGIAYLPVSTAALII--- 113  
 93 ALGRRR-----AWLLIAQCLIVLAICTWAFAD-----PALGSHVYLQVAAAGAVLLGF 139  
 114 --ASQLAFAIFSFVFKHKTFF---TINA-----VLLTVGAALVGMHTETDKP--VH 161  
 140 SAATQDIVDAYRIELAEETQTVLASTYNAGYRIGMIVAGAGALFLAAHLGTAKGNYIY 199  
 162 ETHKQYITGFLITVAAAVMYAFIL-----PLVELAYOKAKQTMSTYTLVLEFQLILCLLA 215  
 200 EAWK---TTVL-TMAAVMLVGIITLLVREPVDRIYKXYSDDYRLVLF-----FIA 250  
 216 SIYSVIGMTIAGDFKQALPK-----EAREFKLGEALFYVVAVFSAILIWOQFFLGAIGLI 269  
 251 VISFVITYYSGQVTEAISKAGNIQDSAAFLGL-EALRELTATGVAIL-VGYLLVKIGVW 308  
 270 ---FSTSSILVS-----GIMISVL-----PTEVLAVIFYHEKFOAEKGL 306  
 309 NQWAKETWAPILDFFKRYGVKALVLLFLIGFFRISDIAGVISNVFYQDLNFSKQI 368  
 307 SLALSWMGFVSIFYGEIKSG 326  
 369 AEAVKYGVIFSLVGGFLGG 388

RESULT 13  
 US-09-328-352-7304  
 ; Sequence 7304, Application US/09328352  
 ; Patent No. 6562958  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Gary L. Breton et al.  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
 ; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: GTC99-03PA  
 ; CURRENT APPLICATION NUMBER: US/09/328,352  
 ; CURRENT FILING DATE: 1999-06-04  
 ; NUMBER OF SEQ ID NOS: 8252  
 ; SEQ ID NO 7304  
 ; LENGTH: 434  
 ; TYPE: PRT  
 ; ORGANISM: Acinetobacter baumannii  
 US-09-328-352-7304

Query Match 6.4%; Score 112; DB 4; Length 434;  
 Best Local Similarity 21.7%; Pred. No. 0.0039;  
 Matches 70; Conservative 56; Mismatches 112; Indels 84; Gaps 19;

12 ILAIG-----NCGGPLIMRLYFNNGGKRIWFST--FLETA-GFPVIFIPILFSYTR 60  
 10 IIALGFMTFALPIGAGNIIFPPVIAQQAQGDHVLAAAGFLITAVGLFVITMAL-----S 64  
 61 RRSNVGDSSTFELIKPRLIIAAVIGILSGFDNYL-----YAYGI-AYL 104  
 65 RMQGSIE-----IISPLGRIASLLITVVC-----YLSVGPLFATPRTATVSYEIGFSYF 115  
 105 PYSTAALIIASQL--AFTAIKSFSE-----MVKKHKTPTTINAVLLTVGAAVL----- 150  
 116 GTSSSSLLIYSALYFAPFVTVWSLYPNKLLDTVGHVLAPKIIAAILGIAAVMIPAGVVP 175



**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:33:00 ; Search time 88 Seconds  
(without alignments)  
1130.190 Million cell updates/sec

Title: US-09-913-767-9  
Perfect score: 1760  
Sequence: 1 MKALVIINCIILAIAGNCG.....QQEESQETQSLSRPISEC 352

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A Geneseq\_29Jan04:\*  
1: Geneseqp1980s:\*  
2: Geneseqp1990s:\*  
3: Geneseqp2000s:\*  
4: Geneseqp2001s:\*  
5: Geneseqp2002s:\*  
6: Geneseqp2003as:\*  
7: Geneseqp2003bs:\*  
8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1744.5	99.1	351	5	ABB91156
2	1128	64.1	356	5	ABB91157
3	1124	63.9	356	3	AAB10684
4	556.5	31.6	367	3	AAG44536
5	540	30.7	178	3	AAG34467
6	522.5	29.7	386	3	AAG16270
7	522.5	29.7	387	3	AAG16269
8	522.5	29.7	394	3	AAG16268
9	521.5	29.6	386	3	AAG52938
10	521.5	29.6	387	3	AAG52937
11	521.5	29.6	394	3	AAG52936
12	516.5	29.3	367	3	AAG22805
13	516.5	29.3	367	3	AAG40302
14	516.5	29.3	382	3	AAG22804
15	516.5	29.3	382	3	AAG40301
16	501.5	28.5	303	3	AAG44537
17	478	27.2	144	3	AAG15103
18	478	27.2	164	3	AAG15102
19	475	27.0	335	3	AAG17233
20	475	27.0	355	3	AAG17232
21	475	27.0	361	3	AAG17231
22	458.5	26.1	358	5	ABB93669
23	433	24.6	250	3	AAG44538
24	320	18.2	107	3	AAG15104
25	194	11.0	134	3	AAG16232

26	151.5	8.6	342	3	AAG51674	Ag51674 Arabidops
27	141.5	8.0	328	3	AAG17291	Ag17291 Arabidops
28	141.5	8.0	400	3	AAG53064	Ag53064 Arabidops
29	141.5	8.0	414	3	AAG53063	Ag53063 Arabidops
30	140.5	8.0	323	3	AAG32021	Ag32021 Arabidops
31	139.5	7.9	286	3	AAG51675	Ag51675 Arabidops
32	137.5	7.8	97	3	AAG16233	Ag16233 Arabidops
33	137.5	7.8	272	3	AAG17292	Ag17292 Arabidops
34	137.5	7.8	277	3	AAG52836	Ag52836 Arabidops
35	137.5	7.8	277	3	AAG23966	Ag23966 Arabidops
36	137.5	7.8	333	3	AAG52835	Ag52835 Arabidops
37	137.5	7.8	333	3	AAG23965	Ag23965 Arabidops
38	137.5	7.8	336	3	AAG23964	Ag23964 Arabidops
39	137.5	7.8	336	3	AAG52834	Ag52834 Arabidops
40	136.5	7.8	269	3	AAG17293	Ag17293 Arabidops
41	136.5	7.8	344	3	AAG53065	Ag53065 Arabidops
42	133.5	7.6	314	6	ABU25229	Abu25229 Protein e
43	133.5	7.6	345	5	ABBA7934	Abba7934 Listeria
44	129	7.3	296	4	AU33602	Au33602 Pseudomon
45	129	7.3	296	6	ABU15685	Abu15685 Protein e

ALIGNMENTS

RESULT 1  
ABB91156  
ID ABB91156 standard; protein; 351 AA.  
XX ABB91156;  
AC ABB91156;  
XX  
DT 31-MAY-2002 (first entry)  
XX  
DE Herbicidally active polypeptide SEQ ID NO 367.  
XX  
KW Herbicidal; plant; agriculture; herbicide.  
XX  
OS Arabidopsis thaliana.  
XX  
PN WO200210210-A2.  
XX  
PD 07-FEB-2002.  
XX  
PF 28-AUG-2001; 2001WO-EP009892.  
XX  
PR 28-AUG-2001; 2001WO-EP009892.  
XX  
PA (FARB ) BAYER AG.  
XX  
PI Tietjen K, Weidler M;  
XX  
DR WPI; 2002-269010/31.  
XX  
PT Identifying plant target proteins for herbicidally active compounds,  
PT comprising aligning and comparing nucleic acid or amino acid sequences  
PT from plant with nucleic acid or amino acid sequences from non-plant  
XX organisms.  
PS Claim 5; SEQ ID NO 367; 261pp + Sequence Listing; English.  
XX  
CC The invention relates to identifying target proteins (ABB90790-ABB94016)  
CC for herbicidally active compounds, comprising aligning and comparing  
CC nucleic acid or amino acid sequences from plant with nucleic acid or  
CC amino acid sequences from non-plant organisms using suitable search  
CC parameters, where plant sequences having an E-value greater by a factor  
CC of 3 than the E-value of most similar non-plant sequences are selected.  
CC The polypeptides or nucleic acids encoding them are useful for  
CC identifying modulators. The identified modulators are useful as  
XX herbicides  
SQ Sequence 351 AA;

Query Match 99.1%; Score 1744.5; DB 5; Length 351;

Best Local Similarity 99.7%; Pred. No. 3.3e-178; Matches 351; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY 1 MVKALVIINCIIILAIKNCGGPLIMRLYFNNGKRIWFSTFLETAGFPVFIPIPLFSYITR 60  
DB 1 MVKALVIINCIIILAIKNCGGPLIMRLYFNNGKRIWFSTFLETAGFPVFIPIPLFSYITR 60  
QY 61 RRSNNVGDSTFFLIKPRLLIAAVIVGILSGFDNYLYAYGIAYLPVSTAALIIASQLAFI 120  
DB 61 RRSNNVGDSTFFLIKPRLLIAAVIVGILSGFDNYLYAYGIAYLPVSTAALIIASQLAFI 120  
QY 121 AIFSFPMVKHKTPTTINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVAAAVM 180  
DB 121 AIFSFPMVKHKTPTTINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVAAAVM 180  
QY 181 YAFILPLVELAYQKAKQTMSTVLVLEFQILICLLASIVSVIGMFIAGDFKALPKAREF 240  
DB 181 YAFILPLVELAYQKAKQTMSTVLVLEFQILICLLASIVSVIGMFIAGDFKALPKAREF 239  
QY 241 KLGEALFYVAVFSAIIWQGFGLGALIGLIFSTSSLVGIMISVLLPITEVLAVIFYHEKF 300  
DB 240 KLGEALFYVAVFSAIIWQGFGLGALIGLIFSTSSLVGIMISVLLPITEVLAVIFYHEKF 299  
QY 301 QAEKGLSLALSLWGFVSFYFYGEIKSGEDKRRIOQESQETQSSLSRPISEC 352  
DB 300 QAEKGLSLALSLWGFVSFYFYGEIKSGEDKRRIOQESQETQSSLSRPISEC 351

RESULT 2  
ABB91157 ID ABB91157 standard; protein; 356 AA.  
XX AC ABB91157;  
DT 31-MAY-2002 (first entry)  
XX DE Herbicidally active polypeptide SEQ ID NO 368.  
XX KW Herbicidal; plant; agriculture; herbicide.  
XX OS Arabidopsis thaliana.  
XX PN WO200210210-A2.  
XX PD 07-FEB-2002.  
XX PF 28-AUG-2001; 2001WO-EP009892.  
XX PR 28-AUG-2001; 2001WO-EP009892.  
XX PA (FARB ) BAYER AG.  
XX PI Tietjen K, Weidler M;  
XX DR WPI; 2002-269010/31.  
XX ID Identifying plant target proteins for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms.  
XX PS Claim 5; SEQ ID NO 368; 261pp + Sequence Listing; English.  
XX CC The invention relates to identifying target proteins (ABB90790-ABB94016) for herbicidally active compounds, comprising aligning and comparing nucleic acid or amino acid sequences from plant with nucleic acid or amino acid sequences from non-plant organisms using suitable search parameters, where plant sequences having an E-value greater by a factor of 3 than the E-value of most similar non-plant sequences are selected. The polypeptides or nucleic acids encoding them are useful for identifying modulators. The identified modulators are useful as herbicides

SQ Sequence 356 AA;  
Query Match 64.1%; Score 1128; DB 5; Length 356;  
Best Local Similarity 60.1%; Pred. NO. 4.5e-112; Indels 12; Gaps 5;  
Matches 215; Conservative 66; Mismatches 65;

QY 1 MVKALVIINCIIILAIKNCGGPLIMRLYFNNGKRIWFSTFLETAGFPVFIPIPLFSYITR 60  
DB 1 MKGGLIIICIIITITCGGPLLTRLYFTNGKRIWFMSFLSTAGFPFIILPLVFLSR 60  
QY 61 RRSNNVGDSTFFLIKPRLLIAAVIVGILSGFDNYLYAYGIAYLPVSTAALIIA 114  
DB 61 RRSNNPNNNAENKRTKLFMETPLFIASIVIGLLTGLDNYLSYGLAYLPVSTSLIIG 120  
QY 115 SOLAFTAIISFFMVKHKTPTTINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLIT 174  
DB 121 TQAFNALFAFLVQKQFTPFSSINAVLLTVGILALHSDGDKPAKESKEVVGFLMT 180  
QY 175 VAAAVMYAFILPLVELAYQKAKQTMSTVTLVLEFQILICLLASIVSVIGMFIAGDFKQALP 234  
DB 181 VVAALLYAFILPLVELTYKARQEIITFPLVLEIQVMCLAATFFCVIGMFIAGDFK-VIA 239  
QY 235 KEAREFKLGEALP-YVNAVFSALIIWQGFGLGALIGLIFSTSSLVGIMISVLLPITEVLAV 293  
DB 240 REAREFKIGSVFYALIVITIGIWOQGFGLGALIGIVFCASSLASGVLISVLLPITEVFAV 299  
QY 294 IFYHEKFOAEKGLSLALSLWGFVSFYFYGEIKSGEDKRRIOQESQETQSSLSRPISE 351  
DB 300 VCFREKFOAEKGVSLLSLWGFVSFYFYGEFKSG--KKVVDKQPQPPETELPIL--PVSD 353

RESULT 3  
AAB10684 ID AAB10684 standard; protein; 356 AA.  
XX AC AAB10684;  
DT 19-JAN-2001 (first entry)  
XX DE A. thaliana PUP1 protein.  
XX KW PUP1; transgenic plant; nucleobase transporter; apical dominance; flowering behaviour; senescence; pesticide distribution.  
XX OS Arabidopsis thaliana.  
XX PN DE19907209-A1.  
XX PD 24-AUG-2000.  
XX PF 19-FEB-1999; 99DE-01007209.  
XX PR 19-FEB-1999; 99DE-01007209.  
XX PA (FROM/) FROMMER W.  
XX PI Gillissen B, Buerkle L, Andre B, Frommer WB;  
XX DR WPI; 2000-565202/53.  
XX DR N-PSDB; AAA97919.  
XX PT Nucleic acid, useful for producing transgenic plants with altered nucleobase transport, encodes a nucleobase transporter protein of Arabidopsis thaliana.  
XX PS Claim 1b; Page 16-17; 24pp; German.  
XX CC This invention describes a novel nucleic acid encoding a plant nucleobase transporter (I). (I) is produced by complementation of a nucleobase transport (NBT)-defective host cell with a plant gene bank by selection of NBT-positive cells. (I) is used to isolate homologous sequences from bacteria, fungi, plants, animals and humans, for expression of the encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting





PR	20-JUL-1999;	99US-0144632P.
PR	20-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0144814P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	21-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	22-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	27-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145919P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	23-AUG-1999;	99US-0150566P.
PR	25-AUG-1999;	99US-0150884P.
PR	26-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	30-AUG-1999;	99US-0151303P.
PR	31-AUG-1999;	99US-0151438P.
PR	01-SEP-1999;	99US-0151930P.
PR	07-SEP-1999;	99US-0152363P.
PR	10-SEP-1999;	99US-0153070P.
PR	13-SEP-1999;	99US-0153758P.
PR	15-SEP-1999;	99US-0154018P.
PR	16-SEP-1999;	99US-0154039P.
PR	20-SEP-1999;	99US-0154779P.
PR	22-SEP-1999;	99US-0155139P.
PR	23-SEP-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159233P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0159584P.
PR	21-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 31.6%; Score 556.5; DB 3; Length 367;		
Best Local Similarity 34.4%; Pred. No. 9.2e-51;		
Matches 115; Conservative 89; Mismatches 121; Indels 9; Gaps 4;		
Qy	5	LVINCIILAIAGNCGGLIMRLYFNNGKRIWFSTFLETAGFPVIFIPILLFSYITRRSN 64
Db	33	LVSVNIFFLIGQQAASVLLGRFYDEGGNSKMWATLVQTAFFILYIPLL----LPSSA 88
Qy	65	NVGDSSTFFELIKPRLLIAAVIVGILSGFDNYLYAYGIAYLPVSTAALIISQALFAIFS 124
Db	89	SVESSESSCSLKIVLI-YVLLGVILAGDNMLYSGVLLKSLASTYSLICATQAFNAVFS 147
Qy	125	FFVVKHKFTPTTINAVLLTVGAALVGMHTETDKPVHETHKQVITGFLITVAAAVYAFI 184
Db	148	YFINAQKFTAILILNSVLLSFAALIALNDADTPSGVSRSKYIVGVFVCTLAASALYSL 207
Qy	185	LPLVELAYQAKQTMSTVTLVLEFOLILCLLASIVSVIGMFIAGDFKQALPKAREFKLGE 244
Db	208	LSLMQFSFKILKRETFSVVLEMQIYTSLVATCVSVIGLFSAGEWR-TLHGEMEGYHKQ 266
Qy	245	ALFYVWAVFSAIIWQGFGLGALIFSTSLVSGIMISVLLPITEVLAVIFYHEKFOAEK 304
Db	267	ASYVLTAVTAVTWQCVSGVGLIFLVTSLSFNSVISTSLAVTPLAALVVRDKMSGVK 326
Qy	305	GLSLALSLWGFVSIFYGEIKSGEDKRIQOESQ 338
Db	327	IMAMLIAWGPASYVY---QNHIDDLKVRQARQQ 357
RESULT 5		
AAG34467		
ID	AAG34467	standard; protein; 178 AA.
XX	AAG34467;	
AC	AAG34467;	
XX	18-OCT-2000	(first entry)
DE	Arabidopsis thaliana	protein fragment SEQ ID NO: 41941.
XX	Arabidopsis thaliana	
KW	Protein identification; signal transduction pathway; metabolic pathway;	
KW	hybridisation assay; genetic mapping; gene expression control; promoter;	
XX	termination sequence.	
OS	Arabidopsis thaliana	
XX	EP1033405-A2.	
PN	06-SEP-2000.	
XX	25-FEB-2000; 2000EP-00301439.	
PF		

XX  
PR 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 29-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 04-MAY-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-01334256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 21-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 08-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 11-AUG-1999; 99US-0148171P.  
PR 12-AUG-1999; 99US-0148319P.  
PR 13-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 16-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151068P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.

Mon Sep 27 08:51:16 2004

PR	29-SEP-1999;	99US-0156596P.	PF	25-FEB-2000;	2000EP-00301439.
PR	04-OCT-1999;	99US-0157117P.	XX	25-FEB-1999;	99US-0121825P.
PR	05-OCT-1999;	99US-0157753P.	PR	05-MAR-1999;	99US-0123180P.
PR	06-OCT-1999;	99US-0157865P.	PR	09-MAR-1999;	99US-0123548P.
PR	07-OCT-1999;	99US-0158029P.	PR	23-MAR-1999;	99US-0125788P.
PR	08-OCT-1999;	99US-0158232P.	PR	25-MAR-1999;	99US-0126264P.
PR	12-OCT-1999;	99US-0158369P.	PR	29-MAR-1999;	99US-0126785P.
PR	13-OCT-1999;	99US-0159293P.	PR	01-APR-1999;	99US-0127462P.
PR	13-OCT-1999;	99US-0159294P.	PR	06-APR-1999;	99US-0128234P.
PR	13-OCT-1999;	99US-0159295P.	PR	08-APR-1999;	99US-0128714P.
PR	14-OCT-1999;	99US-0159323P.	PR	16-APR-1999;	99US-0129845P.
PR	14-OCT-1999;	99US-0159330P.	PR	19-APR-1999;	99US-0130077P.
PR	14-OCT-1999;	99US-0159331P.	PR	21-APR-1999;	99US-0130449P.
PR	14-OCT-1999;	99US-0159637P.	PR	23-APR-1999;	99US-0130510P.
PR	14-OCT-1999;	99US-0159638P.	PR	23-APR-1999;	99US-0130891P.
PR	18-OCT-1999;	99US-0159584P.	PR	28-APR-1999;	99US-0131449P.
PR	21-OCT-1999;	99US-0160741P.	PR	30-APR-1999;	99US-0132048P.
PR	21-OCT-1999;	99US-0160767P.	PR	30-APR-1999;	99US-0132407P.
PR	21-OCT-1999;	99US-0160768P.	PR	04-MAY-1999;	99US-0132484P.
PR	21-OCT-1999;	99US-0160770P.	PR	05-MAY-1999;	99US-0132485P.
PR	21-OCT-1999;	99US-0160814P.	PR	06-MAY-1999;	99US-0132486P.
PR	21-OCT-1999;	99US-0160815P.	PR	06-MAY-1999;	99US-0132487P.
PR	22-OCT-1999;	99US-0160980P.	PR	07-MAY-1999;	99US-0132863P.
PR	22-OCT-1999;	99US-0160981P.	PR	11-MAY-1999;	99US-0132863P.
PR	25-OCT-1999;	99US-0160989P.	PR	14-MAY-1999;	99US-0134218P.
PR	25-OCT-1999;	99US-0161404P.	PR	14-MAY-1999;	99US-0134219P.
PR	25-OCT-1999;	99US-0161405P.	PR	14-MAY-1999;	99US-0134221P.
PR	25-OCT-1999;	99US-0161406P.	PR	14-MAY-1999;	99US-0134370P.
PR	26-OCT-1999;	99US-0161359P.	PR	18-MAY-1999;	99US-0134376P.
PR	26-OCT-1999;	99US-0161360P.	PR	19-MAY-1999;	99US-0134941P.
PR	26-OCT-1999;	99US-0161361P.	PR	20-MAY-1999;	99US-0135124P.
PR	28-OCT-1999;	99US-0161920P.	PR	21-MAY-1999;	99US-0135353P.
PR	28-OCT-1999;	99US-0161922P.	PR	21-MAY-1999;	99US-0135629P.
PR	28-OCT-1999;	99US-0161993P.	PR	24-MAY-1999;	99US-0136021P.
PR	29-OCT-1999;	99US-0162142P.	PR	25-MAY-1999;	99US-0136021P.
Query Match 30.7%; Score 540; DB 3; Length 178;			PR	27-MAY-1999;	99US-0136392P.
Best Local Similarity 60.0%; Pred. No. 2.le-49;			PR	28-MAY-1999;	99US-0136782P.
Matches 108; Conservative 32; Mismatches 34; Indels 6; Gaps 4;			PR	01-JUN-1999;	99US-0137222P.
			PR	03-JUN-1999;	99US-0137528P.
Qy	173	ITVAAMVAFILPLVELAVOKAKQMTSVTLVLEFOLILCLLASIVSVIGMFIAGDFKQA 232	PR	04-JUN-1999;	99US-0137502P.
Db	1	MTVVAALYAFILPLVELYIKKARQITPLVLLEIQMWCLATPCVLGMFIVGDFK-V 59	PR	07-JUN-1999;	99US-0137724P.
Qy	233	LPKEAREFKLGEALF-YVAVFSAIIWQFFLGAIGLIFSTSLVSGIMISVLLPITEVL 291	PR	08-JUN-1999;	99US-0138094P.
Db	60	IMAEAREFKIGSVFYALIVITGIIWQFFLGAIGLIVFCASSLASGVLLISVLLPVEVF 119	PR	10-JUN-1999;	99US-0138540P.
Qy	292	AVIFYHEKFOAEKGLSLALSLMGFVSFYFYGEIKSGEDKRIQEESEQESQSLSRPISE 351	PR	14-JUN-1999;	99US-0139119P.
Db	120	AVVCFREKFOAEKGVLSLLMGFVSFYFYGEFKSG--KKVVDKQPPELPLIL--PVSD 175	PR	16-JUN-1999;	99US-0139452P.
RESULT 6			PR	17-JUN-1999;	99US-0139492P.
AAG16270			PR	18-JUN-1999;	99US-0139454P.
ID	AAG16270 standard; protein; 386 AA.		PR	18-JUN-1999;	99US-0139455P.
AC	AAG16270;		PR	18-JUN-1999;	99US-0139456P.
XX			PR	18-JUN-1999;	99US-0139457P.
DT	17-OCT-2000 (first entry)		PR	18-JUN-1999;	99US-0139458P.
XX			PR	18-JUN-1999;	99US-0139459P.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 16849.		PR	18-JUN-1999;	99US-0139460P.
XX			PR	18-JUN-1999;	99US-0139461P.
XX			PR	18-JUN-1999;	99US-0139462P.
KW	Protein identification; signal transduction pathway; metabolic pathway;		PR	18-JUN-1999;	99US-0139463P.
KW	hybridisation assay; genetic mapping; gene expression control; promoter;		PR	18-JUN-1999;	99US-0139750P.
KW	termination sequence.		PR	21-JUN-1999;	99US-0139763P.
XX	Arabidopsis thaliana.		PR	21-JUN-1999;	99US-0139817P.
OS			PR	22-JUN-1999;	99US-0139899P.
XX			PR	23-JUN-1999;	99US-0140353P.
PN	EP1033405-A2.		PR	23-JUN-1999;	99US-0140354P.
XX			PR	24-JUN-1999;	99US-0140695P.
PD	06-SEP-2000.		PR	28-JUN-1999;	99US-0140823P.
XX			PR	29-JUN-1999;	99US-0140991P.
XX			PR	30-JUN-1999;	99US-0141287P.
XX			PR	01-JUL-1999;	99US-0141842P.
XX			PR	01-JUL-1999;	99US-0142154P.
XX			PR	02-JUL-1999;	99US-0142055P.
XX			PR	06-JUL-1999;	99US-0142390P.
XX			PR	08-JUL-1999;	99US-0142803P.

PR 09-JUL-1999; 99US-0142320P.  
 PR 12-JUL-1999; 99US-0142377P.  
 PR 13-JUL-1999; 99US-0143542P.  
 PR 14-JUL-1999; 99US-0143624P.  
 PR 15-JUL-1999; 99US-0144005P.  
 PR 16-JUL-1999; 99US-0144085P.  
 PR 16-JUL-1999; 99US-0144086P.  
 PR 19-JUL-1999; 99US-0144325P.  
 PR 19-JUL-1999; 99US-0144331P.  
 PR 19-JUL-1999; 99US-0144332P.  
 PR 19-JUL-1999; 99US-0144333P.  
 PR 19-JUL-1999; 99US-0144334P.  
 PR 19-JUL-1999; 99US-0144335P.  
 PR 20-JUL-1999; 99US-0144352P.  
 PR 20-JUL-1999; 99US-0144632P.  
 PR 20-JUL-1999; 99US-0144884P.  
 PR 21-JUL-1999; 99US-0144814P.  
 PR 21-JUL-1999; 99US-0145086P.  
 PR 21-JUL-1999; 99US-0145088P.  
 PR 22-JUL-1999; 99US-0145085P.  
 PR 22-JUL-1999; 99US-0145087P.  
 PR 22-JUL-1999; 99US-0145089P.  
 PR 22-JUL-1999; 99US-0145192P.  
 PR 23-JUL-1999; 99US-0145145P.  
 PR 23-JUL-1999; 99US-0145218P.  
 PR 23-JUL-1999; 99US-0145224P.  
 PR 26-JUL-1999; 99US-0145276P.  
 PR 27-JUL-1999; 99US-0145913P.  
 PR 27-JUL-1999; 99US-0145918P.  
 PR 27-JUL-1999; 99US-0145919P.  
 PR 28-JUL-1999; 99US-0145951P.  
 PR 02-AUG-1999; 99US-0146386P.  
 PR 02-AUG-1999; 99US-0146388P.  
 PR 02-AUG-1999; 99US-0146389P.  
 PR 03-AUG-1999; 99US-0147038P.  
 PR 04-AUG-1999; 99US-0147204P.  
 PR 04-AUG-1999; 99US-0147302P.  
 PR 05-AUG-1999; 99US-0147192P.  
 PR 05-AUG-1999; 99US-0147260P.  
 PR 06-AUG-1999; 99US-0147303P.  
 PR 06-AUG-1999; 99US-0147416P.  
 PR 09-AUG-1999; 99US-0147493P.  
 PR 09-AUG-1999; 99US-0147935P.  
 PR 10-AUG-1999; 99US-0148319P.  
 PR 11-AUG-1999; 99US-0148341P.  
 PR 12-AUG-1999; 99US-0148565P.  
 PR 13-AUG-1999; 99US-0148684P.  
 PR 16-AUG-1999; 99US-0149368P.  
 PR 17-AUG-1999; 99US-0149175P.  
 PR 18-AUG-1999; 99US-0149426P.  
 PR 20-AUG-1999; 99US-0149722P.  
 PR 20-AUG-1999; 99US-0149723P.  
 PR 20-AUG-1999; 99US-0149829P.  
 PR 23-AUG-1999; 99US-0149902P.  
 PR 23-AUG-1999; 99US-0149930P.  
 PR 26-AUG-1999; 99US-0150566P.  
 PR 26-AUG-1999; 99US-0150884P.  
 PR 27-AUG-1999; 99US-0151065P.  
 PR 27-AUG-1999; 99US-0151066P.  
 PR 27-AUG-1999; 99US-0151080P.  
 PR 30-AUG-1999; 99US-0151303P.  
 PR 31-AUG-1999; 99US-0151438P.  
 PR 01-SEP-1999; 99US-0151930P.  
 PR 07-SEP-1999; 99US-0152363P.  
 PR 10-SEP-1999; 99US-0153070P.  
 PR 13-SEP-1999; 99US-0153758P.  
 PR 15-SEP-1999; 99US-0154018P.  
 PR 16-SEP-1999; 99US-0154039P.  
 PR 20-SEP-1999; 99US-0154779P.  
 PR 22-SEP-1999; 99US-0155139P.  
 PR 23-SEP-1999; 99US-0155486P.  
 PR 24-SEP-1999; 99US-0155659P.

PR 28-SEP-1999; 99US-0156458P.  
 PR 29-SEP-1999; 99US-0156596P.  
 PR 04-OCT-1999; 99US-0157117P.  
 PR 05-OCT-1999; 99US-0157753P.  
 PR 06-OCT-1999; 99US-0157865P.  
 PR 07-OCT-1999; 99US-0158029P.  
 PR 08-OCT-1999; 99US-0158232P.  
 PR 12-OCT-1999; 99US-0158369P.  
 PR 13-OCT-1999; 99US-0159293P.  
 PR 13-OCT-1999; 99US-0159294P.  
 PR 13-OCT-1999; 99US-0159295P.  
 PR 14-OCT-1999; 99US-0159329P.  
 PR 14-OCT-1999; 99US-0159330P.  
 PR 14-OCT-1999; 99US-0159331P.  
 PR 14-OCT-1999; 99US-0159637P.  
 PR 18-OCT-1999; 99US-0159584P.  
 PR 21-OCT-1999; 99US-0160741P.  
 PR 21-OCT-1999; 99US-0160767P.  
 PR 21-OCT-1999; 99US-0160768P.  
 PR 21-OCT-1999; 99US-0160770P.  
 PR 21-OCT-1999; 99US-0160814P.  
 PR 21-OCT-1999; 99US-0160815P.  
 PR 22-OCT-1999; 99US-0160980P.  
 PR 22-OCT-1999; 99US-0160981P.  
 PR 22-OCT-1999; 99US-0160989P.  
 PR 25-OCT-1999; 99US-0161404P.  
 PR 25-OCT-1999; 99US-0161405P.  
 PR 25-OCT-1999; 99US-0161406P.  
 PR 26-OCT-1999; 99US-0161359P.  
 PR 26-OCT-1999; 99US-0161360P.  
 PR 26-OCT-1999; 99US-0161361P.  
 PR 28-OCT-1999; 99US-0161920P.  
 PR 28-OCT-1999; 99US-0161992P.  
 PR 28-OCT-1999; 99US-0161993P.  
 PR 29-OCT-1999; 99US-0162142P.

Query Match 29.7%; Score 522.5; DB 3; Length 386;  
 Best Local Similarity 33.8%; Pred. No. 4.3e-47;  
 Matches 116; Conservative 82; Mismatches 126; Indels 19; Gaps 8;

QY 12 ILAIGNCGPLIMRLYPNNGKRIWFSTFETAGFPVIFPLFYSYITRRSN--VGD 68  
 DB 43 LLAGETIATLLGRLYEKGKSTWLETVLQVGF-LTLPVY-YLKPEPSKTKITK 100  
 QY 69 STSFELIKPRLJIAAVIVGILSGFDNYLYAYGTYAVLPVSTAAIJIASOLAFIAFSEFMV 128  
 DB 101 TTSSFL---TLSLVYIGLLVAGHCILYSFGLLYLPVSTFSLISASQAFNAVFSYFLN 157  
 QY 129 KHKFTPTINAVLLTVGAALVGMHTETDKPVHETHK----QYITGFLITVAAAVYAF 183  
 DB 158 SQKITPFIINSVLVLTISSTLLVIQHEPSP-SSTSKSAKSKYVIGYICAVOSSAGYSL 216  
 QY 184 ILPLVELAYOKAKQWMSYTLVLEFQILCLLASIVSIGMFTAGDPKQALPKEARBFKL 243  
 DB 217 VLSLTDYAFEKILKKTFRKAILDMATYPSMWATCVVWVGLFGSGWK-LSTEMEESFOLG 275  
 QY 244 EALFYVAVFSAIWMQFPLGALIGLIFSTSSLVSGIMISVLLPITVLAIVFHEKFAE 303  
 DB 276 KSYILINIGSTISWQACLGISVGLIIEVSSLFSNVISTLCPLFPVPLAVVFRDEMGI 335  
 QY 304 KGLSLALSGLWGFVSYPY---GEIKSGEDKRRIQEESQTEQ 342  
 DB 336 KLIAFLAIWGFVSYGQHYVNDKPEEDQELPQSKKEEEQKQ 378

RESULT 7  
 AAG16269  
 ID AAG16269 standard; protein; 387 AA.  
 XX AAG16269;  
 XX AAG16269;  
 DT 17-OCT-2000 (first entry)

Mon Sep 27 08:51:16 2004

XX Arabidopsis thaliana protein fragment SEQ ID NO: 16848.  
DE Protein identification; signal transduction pathway; metabolic pathway;  
XX hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
KW  
XX  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX  
XX 06-SEP-2000.  
XX  
XX 25-FEB-2000; 2000EP-00301439.  
XX  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.  
PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 28-APR-1999; 99US-0130891P.  
PR 30-APR-1999; 99US-0131449P.  
PR 04-MAY-1999; 99US-0132048P.  
PR 08-MAY-1999; 99US-0132484P.  
PR 14-MAY-1999; 99US-0132485P.  
PR 16-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 18-MAY-1999; 99US-0134219P.  
PR 20-MAY-1999; 99US-0134219P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136382P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.

PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139859P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 20-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145145P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 27-JUL-1999; 99US-0145951P.  
PR 28-JUL-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 06-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 09-AUG-1999; 99US-0148171P.  
PR 10-AUG-1999; 99US-0148319P.  
PR 11-AUG-1999; 99US-0148341P.  
PR 12-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-014926P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 26-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.



PR	17-JUN-1999;	99US-0139492P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139454P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139455P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139456P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139457P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139458P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139459P.	PR	20-AUG-1999;	99US-0149929P.
PR	18-JUN-1999;	99US-0139460P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139461P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139462P.	PR	25-AUG-1999;	99US-0150568P.
PR	18-JUN-1999;	99US-0139463P.	PR	26-AUG-1999;	99US-0150884P.
PR	18-JUN-1999;	99US-0139750P.	PR	27-AUG-1999;	99US-0151065P.
PR	18-JUN-1999;	99US-0139763P.	PR	27-AUG-1999;	99US-0151066P.
PR	21-JUN-1999;	99US-0139817P.	PR	27-AUG-1999;	99US-0151080P.
PR	21-JUN-1999;	99US-0139899P.	PR	30-AUG-1999;	99US-0151303P.
PR	23-JUN-1999;	99US-0140353P.	PR	31-AUG-1999;	99US-0151438P.
PR	23-JUN-1999;	99US-0140354P.	PR	01-SEP-1999;	99US-0151930P.
PR	23-JUN-1999;	99US-0140695P.	PR	07-SEP-1999;	99US-0152363P.
PR	24-JUN-1999;	99US-0140823P.	PR	10-SEP-1999;	99US-0153070P.
PR	28-JUN-1999;	99US-0140991P.	PR	13-SEP-1999;	99US-0153758P.
PR	29-JUN-1999;	99US-0141287P.	PR	15-SEP-1999;	99US-0154018P.
PR	30-JUN-1999;	99US-0141842P.	PR	16-SEP-1999;	99US-0154039P.
PR	01-JUL-1999;	99US-0142154P.	PR	20-SEP-1999;	99US-0154779P.
PR	02-JUL-1999;	99US-0142055P.	PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0142390P.	PR	23-SEP-1999;	99US-0155486P.
PR	08-JUL-1999;	99US-0142803P.	PR	24-SEP-1999;	99US-0155659P.
PR	09-JUL-1999;	99US-0142920P.	PR	28-SEP-1999;	99US-0156458P.
PR	12-JUL-1999;	99US-0142977P.	PR	29-SEP-1999;	99US-0156596P.
PR	13-JUL-1999;	99US-0143542P.	PR	04-OCT-1999;	99US-0157117P.
PR	14-JUL-1999;	99US-0143624P.	PR	05-OCT-1999;	99US-0157753P.
PR	15-JUL-1999;	99US-0144005P.	PR	06-OCT-1999;	99US-0157865P.
PR	16-JUL-1999;	99US-0144086P.	PR	07-OCT-1999;	99US-0158029P.
PR	16-JUL-1999;	99US-0144086P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144331P.	PR	12-OCT-1999;	99US-0158369P.
PR	19-JUL-1999;	99US-0144332P.	PR	12-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144334P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144352P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144632P.	PR	14-OCT-1999;	99US-0159637P.
PR	20-JUL-1999;	99US-0144884P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0144814P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145086P.	PR	18-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-0145085P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-0145087P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-0145189P.	PR	21-OCT-1999;	99US-0160815P.
PR	22-JUL-1999;	99US-0145192P.	PR	22-OCT-1999;	99US-0160980P.
PR	22-JUL-1999;	99US-0145145P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-0145224P.	PR	25-OCT-1999;	99US-0161404P.
PR	26-JUL-1999;	99US-0145276P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-0145913P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-0145918P.	PR	25-OCT-1999;	99US-0161359P.
PR	27-JUL-1999;	99US-0145919P.	PR	26-OCT-1999;	99US-0161360P.
PR	28-JUL-1999;	99US-0145951P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-0146386P.	PR	26-OCT-1999;	99US-0161920P.
PR	02-AUG-1999;	99US-0146388P.	PR	28-OCT-1999;	99US-0161922P.
PR	02-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161992P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161993P.
PR	04-AUG-1999;	99US-0147204P.	PR	29-OCT-1999;	99US-0162142P.
PR	04-AUG-1999;	99US-0147302P.			
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148315P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
			Query Match		29.7%; Score 522.5; DB 3; Length 394;
			Best Local Similarity		33.8%; Pred. No. 4.5e-47;
			Matches 116; Conservative		82; Mismatches 126; Indels 19; Gaps 8;
Qy	12	ILAINCGGPLIMRYFNNGKRIFWSTFLETAGPVIFIPILFSYITRRSNN---	Qy	12	ILAINCGGPLIMRYFNNGKRIFWSTFLETAGPVIFIPILFSYITRRSNN---
Db	51	LLLAGETIATLGRUYEKGKSTWLETIVOLVGPP-LTLPQY-YLKPEPSKTIYTKK	Db	51	LLLAGETIATLGRUYEKGKSTWLETIVOLVGPP-LTLPQY-YLKPEPSKTIYTKK
Qy	69	STSFLLIKPRLLIAAVIVIGILSGFDNYLYAYGIAVLPVSTALIIASOLAFIAIFSF	Qy	69	STSFLLIKPRLLIAAVIVIGILSGFDNYLYAYGIAVLPVSTALIIASOLAFIAIFSF



```
Db 109 TTSSFL---TSLVYIGLGLLVAGHCILYSGFLLYLPVSTFSLISASQLAFNAVFSYFLN 165
Qy 129 KHRFTPTINAVLLTVGAALVGLMHTETDKPVHETHK-----QYITGLITVAAAVMYAF 183
Db 166 SQKITPILNSLVLLTISLLVQHEPESP-SSTSKSAKSKYVIGYICAVGSSAGYSL 224
Qy 184 ILPLVELAYOKAQWTSYTLVLBFQLILCLLASIVSIGFIAGDFKQALPKKAREPKLG 243
Db 225 VLSLTDYAFKILKXKTKFKAILDMATYPSMWATCVVVGFLFGSGGKK-LSTEMEERQLG 283
Qy 244 EALFYVAVFSAIIWQGFLLGAIGLIPSTSSVSGIMISVLLPITEVLAVIFYHEKFAE 303
Db 284 KSSYLINIGSTISWQACILGSLVGLIIEVSLFSNVISTICLPVPLVAVFFRDEMGI 343
Qy 304 KGLSLALSLMGFSYFYF---GRIKSGEDKRRIQEESQTEQ 342
Db 344 KLIAFLAINGFVSYGYQHYVNDKRPEDQELPQSKKEERQK 386

RESULT 9
AAG52938
ID AAG52938 standard; protein; 386 AA.
XX AC AAG52938;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 67347.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 28-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 30-APR-1999; 99US-0132407P.
XX PR 04-MAY-1999; 99US-0132484P.
XX PR 05-MAY-1999; 99US-0132485P.
XX PR 06-MAY-1999; 99US-0132486P.
XX PR 07-MAY-1999; 99US-0132487P.
XX PR 07-MAY-1999; 99US-0132863P.
XX PR 11-MAY-1999; 99US-0134256P.
XX PR 14-MAY-1999; 99US-0134218P.
XX PR 14-MAY-1999; 99US-0134219P.
XX PR 14-MAY-1999; 99US-0134221P.
XX PR 14-MAY-1999; 99US-0134370P.
XX PR 18-MAY-1999; 99US-0134768P.
XX PR 19-MAY-1999; 99US-0134941P.
XX PR 20-MAY-1999; 99US-0135124P.
XX PR 21-MAY-1999; 99US-0135353P.
XX PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142377P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 18-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144884P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 23-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
```

```
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 06-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147933P.
PR 09-AUG-1999; 99US-0148171P.
PR 10-AUG-1999; 99US-0148319P.
PR 11-AUG-1999; 99US-0148341P.
PR 12-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 13-AUG-1999; 99US-0149368P.
PR 16-AUG-1999; 99US-0149175P.
PR 17-AUG-1999; 99US-0149426P.
PR 18-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 20-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 27-AUG-1999; 99US-0151303P.
PR 30-AUG-1999; 99US-0151438P.
PR 31-AUG-1999; 99US-0151930P.
PR 01-SEP-1999; 99US-0152363P.
PR 07-SEP-1999; 99US-0153070P.
PR 10-SEP-1999; 99US-0153758P.
PR 13-SEP-1999; 99US-0154018P.
PR 15-SEP-1999; 99US-0154039P.
PR 16-SEP-1999; 99US-0154779P.
PR 20-SEP-1999; 99US-0155139P.
PR 22-SEP-1999; 99US-0155486P.
PR 23-SEP-1999; 99US-0155659P.
PR 24-SEP-1999; 99US-0156458P.
PR 28-SEP-1999; 99US-0156596P.
PR 29-SEP-1999; 99US-0157117P.
PR 04-OCT-1999; 99US-0157573P.
PR 05-OCT-1999; 99US-0157865P.
PR 06-OCT-1999; 99US-0158029P.
PR 07-OCT-1999; 99US-0158232P.
PR 08-OCT-1999; 99US-0158369P.
PR 12-OCT-1999; 99US-0159233P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 13-OCT-1999; 99US-0159329P.
PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161992P.

PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match 29.6%; Score 521.5; DB 3; Length 386;
Best Local Similarity 33.8%; Pred. No. 5.6e-47;
Matches 116; Conservative 82; Mismatches 126; Indels 19; Gaps 8;

QY 12 ILAIGNCGPLIMRLYFNNGKRIWFSTFLETAGFPVIFIPLLFSYITRRSNN---VGD 68
DB 43 LLLAGETIATLLGRLLYKGGKSTWLETVLVOLVGFP-LTLPCYY-YLKPEPSKTKITTK 100
QY 69 STSFFELIKPRLIIAAVYIGILSGFDNYLYAGIAYLPVSTAALIISOLAFIAIFSPFMV 128
DB 101 TTSSFL---TSLVVYIGLLVAGHCILYSGLLYLPVSTFSLISASQAFNAFVSIFLN 157
QY 129 KHKFTPTTNAVLLTVGAANVLGMHETDKPVHETHK-----QYITGLITVAAMVYAF 183
DB 158 SQKITPFIILNSLVLTISTLLVIOHEPESP-SSTSKSAASKYVIGYICAVGSSAGYSL 216
QY 184 ILPLVELAYOKAKQTMSTYTLVLEFQLIILCLLASIVSVIGMFTAGDFKQALPKBAEFKLG 243
DB 217 VLSTDYAFEKILKXKTFKAILDMATYPSMWATCVVVVGLFGSGGWKK-LSTEMEEFQLG 275
QY 244 EALFYVAVFSAIIWQGFPLGAIGLIFSTSSLVSGMISVLLPITEVLAVIFYHEKFOAE 303
DB 276 KSSVILINIGSTISMQACLIGSVGLIIEVSSLFSNVISTLCPLVPVPLAVVFRDEMGI 335
QY 304 KGLSLALSGLWGFVSFYF---GEIKSGEDKRIQOESEQTEQ 342
DB 336 KLVAMELAINGFVSIGYQHYVNDRKPEEDQQLPQSKEEBEQK 378

RESULT 10
AAGS2937
ID AAGS2937 standard; protein; 387 AA.
XX AC AAGS2937;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 67346.
XX KW Protein identification; signal transduction pathway; metabolic pathway;
XX KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX KW termination sequence.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
XX PR 05-MAR-1999; 99US-0123180P.
XX PR 09-MAR-1999; 99US-0123548P.
XX PR 23-MAR-1999; 99US-0125788P.
XX PR 25-MAR-1999; 99US-0126264P.
XX PR 29-MAR-1999; 99US-0126785P.
XX PR 01-APR-1999; 99US-0127462P.
XX PR 06-APR-1999; 99US-0128234P.
XX PR 08-APR-1999; 99US-0128714P.
XX PR 16-APR-1999; 99US-0129845P.
XX PR 19-APR-1999; 99US-0130077P.
XX PR 21-APR-1999; 99US-0130449P.
XX PR 23-APR-1999; 99US-0130510P.
XX PR 23-APR-1999; 99US-0130891P.
XX PR 28-APR-1999; 99US-0131449P.
XX PR 30-APR-1999; 99US-0132048P.
XX PR 30-APR-1999; 99US-0132407P.
XX PR 04-MAY-1999; 99US-0132484P.
XX PR 05-MAY-1999; 99US-0132485P.
```

PR 06-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135829P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 21-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143342P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144321P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144684P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0132486P.  
PR 22-JUL-1999; 99US-0132487P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145193P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 27-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 23-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 14-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.



PR	19-JUL-1999;	99US-0144331P.	PR	13-OCT-1999;	99US-0159293P.
PR	19-JUL-1999;	99US-0144332P.	PR	13-OCT-1999;	99US-0159294P.
PR	19-JUL-1999;	99US-0144333P.	PR	13-OCT-1999;	99US-0159295P.
PR	19-JUL-1999;	99US-0144334P.	PR	14-OCT-1999;	99US-0159329P.
PR	19-JUL-1999;	99US-0144335P.	PR	14-OCT-1999;	99US-0159330P.
PR	20-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159331P.
PR	20-JUL-1999;	99US-0144332P.	PR	14-OCT-1999;	99US-0159637P.
PR	21-JUL-1999;	99US-0144814P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-01445086P.	PR	21-OCT-1999;	99US-0160767P.
PR	21-JUL-1999;	99US-01445088P.	PR	21-OCT-1999;	99US-0160768P.
PR	22-JUL-1999;	99US-01445087P.	PR	21-OCT-1999;	99US-0160770P.
PR	22-JUL-1999;	99US-01445089P.	PR	21-OCT-1999;	99US-0160814P.
PR	22-JUL-1999;	99US-01445192P.	PR	21-OCT-1999;	99US-0160815P.
PR	23-JUL-1999;	99US-01445145P.	PR	22-OCT-1999;	99US-0160980P.
PR	23-JUL-1999;	99US-01445218P.	PR	22-OCT-1999;	99US-0160981P.
PR	23-JUL-1999;	99US-01445224P.	PR	22-OCT-1999;	99US-0160989P.
PR	26-JUL-1999;	99US-01445276P.	PR	25-OCT-1999;	99US-0161404P.
PR	27-JUL-1999;	99US-01445913P.	PR	25-OCT-1999;	99US-0161405P.
PR	27-JUL-1999;	99US-01445918P.	PR	25-OCT-1999;	99US-0161406P.
PR	27-JUL-1999;	99US-01445919P.	PR	26-OCT-1999;	99US-0161359P.
PR	28-JUL-1999;	99US-01445951P.	PR	26-OCT-1999;	99US-0161360P.
PR	02-AUG-1999;	99US-01446386P.	PR	26-OCT-1999;	99US-0161361P.
PR	02-AUG-1999;	99US-01446388P.	PR	28-OCT-1999;	99US-0161362P.
PR	02-AUG-1999;	99US-01446389P.	PR	28-OCT-1999;	99US-0161920P.
PR	03-AUG-1999;	99US-0147038P.	PR	28-OCT-1999;	99US-0161922P.
PR	04-AUG-1999;	99US-0147204P.	PR	28-OCT-1999;	99US-0161933P.
PR	04-AUG-1999;	99US-0147302P.	PR	29-OCT-1999;	99US-0162142P.
PR	05-AUG-1999;	99US-0147192P.			
PR	05-AUG-1999;	99US-0147260P.			
PR	06-AUG-1999;	99US-0147303P.			
PR	06-AUG-1999;	99US-0147416P.			
PR	09-AUG-1999;	99US-0147493P.			
PR	09-AUG-1999;	99US-0147935P.			
PR	10-AUG-1999;	99US-0148171P.			
PR	11-AUG-1999;	99US-0148319P.			
PR	12-AUG-1999;	99US-0148341P.			
PR	13-AUG-1999;	99US-0148565P.			
PR	13-AUG-1999;	99US-0148684P.			
PR	16-AUG-1999;	99US-0149368P.			
PR	17-AUG-1999;	99US-0149175P.			
PR	18-AUG-1999;	99US-0149426P.			
PR	20-AUG-1999;	99US-0149722P.			
PR	20-AUG-1999;	99US-0149723P.			
PR	20-AUG-1999;	99US-0149929P.			
PR	23-AUG-1999;	99US-0149802P.			
PR	23-AUG-1999;	99US-0149930P.			
PR	25-AUG-1999;	99US-0150566P.			
PR	26-AUG-1999;	99US-0150884P.			
PR	27-AUG-1999;	99US-0151065P.			
PR	27-AUG-1999;	99US-0151066P.			
PR	27-AUG-1999;	99US-0151080P.			
PR	30-AUG-1999;	99US-0151303P.			
PR	31-AUG-1999;	99US-0151438P.			
PR	01-SEP-1999;	99US-0151930P.			
PR	07-SEP-1999;	99US-0152363P.			
PR	10-SEP-1999;	99US-0153070P.			
PR	13-SEP-1999;	99US-0153758P.			
PR	15-SEP-1999;	99US-0154018P.			
PR	16-SEP-1999;	99US-0154039P.			
PR	20-SEP-1999;	99US-0154779P.			
PR	22-SEP-1999;	99US-0155139P.			
PR	23-SEP-1999;	99US-0155486P.			
PR	24-SEP-1999;	99US-0155659P.			
PR	28-SEP-1999;	99US-0156458P.			
PR	29-SEP-1999;	99US-0156596P.			
PR	04-OCT-1999;	99US-0157117P.			
PR	05-OCT-1999;	99US-0157753P.			
PR	06-OCT-1999;	99US-0157865P.			
PR	07-OCT-1999;	99US-0158029P.			
PR	08-OCT-1999;	99US-0158232P.			
PR	12-OCT-1999;	99US-0158369P.			

Query Match 29.8%; Score 521.5; DB 3; Length 394;

Best Local Similarity 33.8%; Pred. No. 5.7e-47;

Matches 116; Conservative 82; Mismatches 126; Indels 19; Gaps 8;

QY 12 ILAIGNCGGLINRLYFNNGGKRIWFSTLETAGFPVIFPLPFSYITRRSNN---VGD 68

Db 51 LLAGETIATLLGRLYERGGKSTWLETVLQLVGFP-LTLPCYY-YLKPEPSTKITTKK 108

QY 69 STSFLLIKPRLIAAVIVGILSGFDNYLYAYGIAYLPVSTAALIIASQLAFIAIFSPFMV 128

Db 109 TTSSFL--TSLVYIGLGLLVAGHCILYSPGLLYLPVSTFSLISASQLAFNAVFSYFLN 165

QY 129 KHKFTPTINAVLLTVGAALVGMHTETDKPVHETHK-----QYTGFLITVAAVWYAF 183

Db 166 SQKITPFIILSLVLLTISSTLLVIQHEPESP-SSTSKSAKSKYVIGYICAVGSAGYSL 224

QY 184 ILPLVELAYQKAKQMTSYTLVLEFQLLCLLASIVSVIGMFIAGDFKQALPKAREKLG 243

Db 225 VLSTDYAFEKILDKYTFKAILDMATYPSMWATCVVVVWGLFGSGGWKK-LSTEMEBFQLG 283

QY 244 EALFYVVAVPSAIWQFFLGALIGFSTSSLVGIMISVLLPITEVLAVIFYHEKFOAE 303

Db 284 KSSYILINIGSTISWQACILGSLVGLIIEVSSLFNSVITSLCLPVPVLAFFRDEMGI 343

QY 304 KGLSLALSLWGFVSYPY---GEIKSGEDKRRIQBESQETE 342

Db 344 KLVAMFLAIWGFVSYGQYHVNDKPEEDQELPQSKEEBEQK 386

RESULT 12

AAG22805

ID AAG22805 standard; protein; 367 AA.

XX

AC AAG22805;

XX

DT 17-OCT-2000 (first entry)

XX

Arabidopsis thaliana protein fragment SEQ ID NO: 25874.

XX

Protein identification; signal transduction pathway; metabolic pathway;

hybridisation assay; genetic mapping; gene expression control; promoter;

termination sequence.

XX

OS Arabidopsis thaliana.

us-09-913-767-9.rag

Mon Sep 27 08:51:16 2004

XX	PR	01-JUL-1999;	99US-0141842P.
PN	PR	01-JUL-1999;	99US-0142154P.
XX	PR	02-JUL-1999;	99US-0142055P.
PN	PR	06-JUL-1999;	99US-0142390P.
PD	PR	08-JUL-1999;	99US-0142803P.
XX	PR	09-JUL-1999;	99US-0142920P.
PF	PR	12-JUL-1999;	99US-0142977P.
XX	PR	13-JUL-1999;	99US-0143542P.
XX	PR	14-JUL-1999;	99US-0143624P.
PR	PR	15-JUL-1999;	99US-0144005P.
PR	PR	16-JUL-1999;	99US-0144085P.
PR	PR	16-JUL-1999;	99US-0144086P.
PR	PR	19-JUL-1999;	99US-0144325P.
PR	PR	19-JUL-1999;	99US-0144331P.
PR	PR	19-JUL-1999;	99US-0144332P.
PR	PR	19-JUL-1999;	99US-0144333P.
PR	PR	19-JUL-1999;	99US-0144334P.
PR	PR	19-JUL-1999;	99US-0144335P.
PR	PR	20-JUL-1999;	99US-0144352P.
PR	PR	20-JUL-1999;	99US-0144632P.
PR	PR	20-JUL-1999;	99US-0144884P.
PR	PR	21-JUL-1999;	99US-0144814P.
PR	PR	21-JUL-1999;	99US-0145086P.
PR	PR	21-JUL-1999;	99US-0145088P.
PR	PR	22-JUL-1999;	99US-0145085P.
PR	PR	22-JUL-1999;	99US-0145087P.
PR	PR	22-JUL-1999;	99US-0145089P.
PR	PR	22-JUL-1999;	99US-0145192P.
PR	PR	22-JUL-1999;	99US-0145145P.
PR	PR	23-JUL-1999;	99US-0145218P.
PR	PR	23-JUL-1999;	99US-0145224P.
PR	PR	26-JUL-1999;	99US-0145276P.
PR	PR	27-JUL-1999;	99US-0145913P.
PR	PR	27-JUL-1999;	99US-0145918P.
PR	PR	27-JUL-1999;	99US-0145919P.
PR	PR	28-JUL-1999;	99US-0145951P.
PR	PR	02-AUG-1999;	99US-0146386P.
PR	PR	02-AUG-1999;	99US-0146388P.
PR	PR	02-AUG-1999;	99US-0146389P.
PR	PR	03-AUG-1999;	99US-0147038P.
PR	PR	04-AUG-1999;	99US-0147204P.
PR	PR	04-AUG-1999;	99US-0147302P.
PR	PR	05-AUG-1999;	99US-0147192P.
PR	PR	05-AUG-1999;	99US-0147260P.
PR	PR	06-AUG-1999;	99US-0147303P.
PR	PR	06-AUG-1999;	99US-0147416P.
PR	PR	09-AUG-1999;	99US-0147493P.
PR	PR	09-AUG-1999;	99US-0147935P.
PR	PR	10-AUG-1999;	99US-0148171P.
PR	PR	11-AUG-1999;	99US-0148319P.
PR	PR	12-AUG-1999;	99US-0148341P.
PR	PR	13-AUG-1999;	99US-0148565P.
PR	PR	13-AUG-1999;	99US-0148684P.
PR	PR	16-AUG-1999;	99US-0149368P.
PR	PR	17-AUG-1999;	99US-0149175P.
PR	PR	18-AUG-1999;	99US-0149436P.
PR	PR	20-AUG-1999;	99US-0149722P.
PR	PR	20-AUG-1999;	99US-0149723P.
PR	PR	20-AUG-1999;	99US-0149929P.
PR	PR	23-AUG-1999;	99US-0149902P.
PR	PR	23-AUG-1999;	99US-0149930P.
PR	PR	25-AUG-1999;	99US-0150566P.
PR	PR	26-AUG-1999;	99US-0150884P.
PR	PR	27-AUG-1999;	99US-0151065P.
PR	PR	27-AUG-1999;	99US-0151066P.
PR	PR	27-AUG-1999;	99US-0151080P.
PR	PR	30-AUG-1999;	99US-0151303P.
PR	PR	31-AUG-1999;	99US-0151438P.
PR	PR	01-SEP-1999;	99US-0151930P.
PR	PR	07-SEP-1999;	99US-0152363P.
PR	PR	10-SEP-1999;	99US-0153070P.
PR	PR	13-SEP-1999;	99US-0153758P.
PR	PR	15-SEP-1999;	99US-0154018P.



20-AUG-1999; 99US-0149929P.  
23-AUG-1999; 99US-0149902P.





PR	09-AUG-1999;	99US-0147935P.	Db	16	MVKPVRKSLTLIVTYFFELFGSIASSLLAKYFYVYGGSSRWVSTWVQAGPPLLLILI	75
PR	10-AUG-1999;	99US-0148171P.				
PR	11-AUG-1999;	99US-0148319P.	Qy	54	LFYS-----ITRRNNVGDSTSPFLIKPRLLIAAVIVGILSGPDNYLYAYGIAYLPVSTA	109
PR	12-AUG-1999;	99US-0148341P.				
PR	13-AUG-1999;	99US-0148565P.	Db	76	YPHYVLTXTTRR-----PPTRETL---RHLIFSVLGLVGLFNNFLFSWGTSLVPVSTS	127
PR	13-AUG-1999;	99US-0148684P.				
PR	16-AUG-1999;	99US-0149368P.	Qy	110	ALIITASOLAFIAFSFFMWKHTPFTINAVVLLTVGAALVGMHTETDKPVHETHKOYIT	169
PR	17-AUG-1999;	99US-0149173P.				
PR	18-AUG-1999;	99US-0149428P.	Db	128	SLLSSTQLVFTLLISRIIVKOKITFNSLNCVLLTSSVLLALDSSKDKPSGLTKTKYFI	187
PR	20-AUG-1999;	99US-0149723P.				
PR	20-AUG-1999;	99US-0149723P.	Qy	170	GELITVAAAAMVAFILPLVELAYQAKQTMGYTLVLEFQLILCLLASIVSIVGMFIAGDF	229
PR	20-AUG-1999;	99US-0149023P.	Db	188	GYVSTIGAGLLFALYLPVTEKLY---RTVICYAMVMEVQLVMEFAATVFATIGACGGP	244
PR	23-AUG-1999;	99US-0149930P.				
PR	23-AUG-1999;	99US-0150566P.	Qy	230	KQALPKAREFKLGEALFYVAVSAII-WQGFLLGAIGLIFTSLSLVSGIMISVLLPIT	288
PR	25-AUG-1999;	99US-0150884P.				
PR	26-AUG-1999;	99US-0151065P.	Db	245	KEMVKEANHVTKGTFTYTFAILANVVVWQLSFAATSGMYLTSGITGGICMTALLMN	304
PR	27-AUG-1999;	99US-0151066P.				
PR	27-AUG-1999;	99US-0151080P.	Qy	289	EVLAVIFVHEKFOAEKGLSLALSLMGFVSFYFGEIKSGEDKRRIOQESQETQSSSL	345
PR	30-AUG-1999;	99US-0151303P.	Db	305	VIGGVVAYGDVFGGVKIVSTVLCINGFSSTYTGMY-----MMKKKEKKEYSVGV	355
PR	31-AUG-1999;	99US-0151438P.				
PR	01-SEP-1999;	99US-0151930P.				
PR	07-SEP-1999;	99US-0152363P.				
PR	10-SEP-1999;	99US-0153070P.				
PR	13-SEP-1999;	99US-0153758P.				
PR	15-SEP-1999;	99US-0154018P.				
PR	16-SEP-1999;	99US-0154039P.				
PR	20-SEP-1999;	99US-0154779P.				
PR	22-SEP-1999;	99US-0155139P.				
PR	23-SEP-1999;	99US-0155486P.				
PR	24-SEP-1999;	99US-0155659P.				
PR	28-SEP-1999;	99US-0156458P.				
PR	29-SEP-1999;	99US-0156596P.				
PR	04-OCT-1999;	99US-0157117P.				
PR	05-OCT-1999;	99US-0157753P.				
PR	06-OCT-1999;	99US-0157865P.				
PR	07-OCT-1999;	99US-0158029P.				
PR	08-OCT-1999;	99US-0158232P.				
PR	12-OCT-1999;	99US-0158369P.				
PR	13-OCT-1999;	99US-0159291P.				
PR	13-OCT-1999;	99US-0159295P.				
PR	14-OCT-1999;	99US-0159339P.				
PR	14-OCT-1999;	99US-0159331P.				
PR	14-OCT-1999;	99US-0159637P.				
PR	14-OCT-1999;	99US-0159638P.				
PR	18-OCT-1999;	99US-0159584P.				
PR	21-OCT-1999;	99US-0160741P.				
PR	21-OCT-1999;	99US-0160767P.				
PR	21-OCT-1999;	99US-0160768P.				
PR	21-OCT-1999;	99US-0160770P.				
PR	21-OCT-1999;	99US-0160814P.				
PR	21-OCT-1999;	99US-0160815P.				
PR	22-OCT-1999;	99US-0160980P.				
PR	22-OCT-1999;	99US-0160981P.				
PR	22-OCT-1999;	99US-0160989P.				
PR	25-OCT-1999;	99US-0161404P.				
PR	25-OCT-1999;	99US-0161405P.				
PR	25-OCT-1999;	99US-0161359P.				
PR	26-OCT-1999;	99US-0161360P.				
PR	26-OCT-1999;	99US-0161361P.				
PR	26-OCT-1999;	99US-0161920P.				
PR	28-OCT-1999;	99US-0161992P.				
PR	28-OCT-1999;	99US-0161993P.				
PR	29-OCT-1999;	99US-0162142P.				
Query Match 29.3%; Score 516.5; DB 3; Length 382;						
Best Local Similarity 32.2%; Pred. No. 1.9e-46;						
Matches 115; Conservative 81; Mismatches 132; Indels 29; Gaps 7;						
1 MKV-----ALVIINCIILAIAGCGPLIMRLYFNNGKRIWFSTFLETAGFPVIFPL 53						
Qy						

RESULT 15

AAG40301

ID AAG40301 standard; protein; 382 AA.

XX AC AAG40301;

XX DT 18-OCT-2000 (first entry)

XX DE Arabidopsis thaliana protein fragment SEQ ID NO: 49986.

XX KW Protein identification; signal transduction pathway; metabolic pathway; hybridisation assay; genetic mapping; gene expression control; promoter; termination sequence.

XX OS Arabidopsis thaliana.

XX PN EP1033405-A2.

XX PD 06-SEP-2000.

XX PF 25-FEB-2000; 2000EP-00301439.

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 29-MAR-1999; 99US-0126264P.

PR 01-APR-1999; 99US-0126785P.

PR 06-APR-1999; 99US-0127462P.

PR 08-APR-1999; 99US-0128234P.

PR 16-APR-1999; 99US-0128714P.

PR 19-APR-1999; 99US-0129845P.

PR 21-APR-1999; 99US-0130077P.

PR 23-APR-1999; 99US-0130449P.

PR 23-APR-1999; 99US-0130510P.

PR 28-APR-1999; 99US-0130891P.

PR 30-APR-1999; 99US-0131449P.

PR 30-APR-1999; 99US-0132048P.

PR 04-MAY-1999; 99US-0132407P.

PR 05-MAY-1999; 99US-0132484P.

PR 06-MAY-1999; 99US-0132485P.

PR 07-MAY-1999; 99US-0132487P.

PR 11-MAY-1999; 99US-0132863P.

PR 14-MAY-1999; 99US-0134256P.

PR 14-MAY-1999; 99US-0134218P.

PR 14-MAY-1999; 99US-0134219P.

PR 14-MAY-1999; 99US-0134221P.

PR 14-MAY-1999; 99US-0134370P.

PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 22-MAY-1999; 99US-0135629P.  
PR 23-MAY-1999; 99US-0136021P.  
PR 24-MAY-1999; 99US-0136382P.  
PR 25-MAY-1999; 99US-0136782P.  
PR 26-MAY-1999; 99US-0137222P.  
PR 27-MAY-1999; 99US-0137528P.  
PR 28-MAY-1999; 99US-0137750P.  
PR 01-JUN-1999; 99US-0137724P.  
PR 02-JUN-1999; 99US-0138094P.  
PR 03-JUN-1999; 99US-0138540P.  
PR 04-JUN-1999; 99US-0138847P.  
PR 05-JUN-1999; 99US-0139119P.  
PR 06-JUN-1999; 99US-0139452P.  
PR 07-JUN-1999; 99US-0139453P.  
PR 08-JUN-1999; 99US-0139457P.  
PR 09-JUN-1999; 99US-0139458P.  
PR 10-JUN-1999; 99US-0139459P.  
PR 11-JUN-1999; 99US-0139460P.  
PR 12-JUN-1999; 99US-0139461P.  
PR 13-JUN-1999; 99US-0139462P.  
PR 14-JUN-1999; 99US-0139463P.  
PR 15-JUN-1999; 99US-0139750P.  
PR 16-JUN-1999; 99US-0139763P.  
PR 17-JUN-1999; 99US-0139763P.  
PR 18-JUN-1999; 99US-0139817P.  
PR 19-JUN-1999; 99US-0139859P.  
PR 20-JUN-1999; 99US-0140353P.  
PR 21-JUN-1999; 99US-0140354P.  
PR 22-JUN-1999; 99US-0140695P.  
PR 23-JUN-1999; 99US-0140823P.  
PR 24-JUN-1999; 99US-0140991P.  
PR 25-JUN-1999; 99US-0141287P.  
PR 26-JUN-1999; 99US-0141842P.  
PR 27-JUN-1999; 99US-0142154P.  
PR 28-JUN-1999; 99US-0142055P.  
PR 29-JUN-1999; 99US-0142390P.  
PR 30-JUN-1999; 99US-0142803P.  
PR 01-JUL-1999; 99US-0142920P.  
PR 02-JUL-1999; 99US-0142977P.  
PR 03-JUL-1999; 99US-0143542P.  
PR 04-JUL-1999; 99US-0143624P.  
PR 05-JUL-1999; 99US-0144005P.  
PR 06-JUL-1999; 99US-0144085P.  
PR 07-JUL-1999; 99US-0144086P.  
PR 08-JUL-1999; 99US-0144325P.  
PR 09-JUL-1999; 99US-0144331P.  
PR 10-JUL-1999; 99US-0144332P.  
PR 11-JUL-1999; 99US-0144333P.  
PR 12-JUL-1999; 99US-0144334P.  
PR 13-JUL-1999; 99US-0144335P.  
PR 14-JUL-1999; 99US-0144352P.  
PR 15-JUL-1999; 99US-0144632P.  
PR 16-JUL-1999; 99US-0144884P.  
PR 17-JUL-1999; 99US-0144884P.  
PR 18-JUL-1999; 99US-0145086P.  
PR 19-JUL-1999; 99US-0145088P.  
PR 20-JUL-1999; 99US-0145085P.  
PR 21-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 24-JUL-1999; 99US-0145218P.  
PR 25-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 28-JUL-1999; 99US-0145918P.

PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 29-JUL-1999; 99US-0146386P.  
PR 30-JUL-1999; 99US-0146388P.  
PR 31-JUL-1999; 99US-0146389P.  
PR 01-AUG-1999; 99US-0146389P.  
PR 02-AUG-1999; 99US-0147038P.  
PR 03-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 07-AUG-1999; 99US-0147303P.  
PR 08-AUG-1999; 99US-0147303P.  
PR 09-AUG-1999; 99US-0147303P.  
PR 10-AUG-1999; 99US-0147303P.  
PR 11-AUG-1999; 99US-0147303P.  
PR 12-AUG-1999; 99US-0147303P.  
PR 13-AUG-1999; 99US-0147303P.  
PR 14-AUG-1999; 99US-0147303P.  
PR 15-AUG-1999; 99US-0147303P.  
PR 16-AUG-1999; 99US-0147303P.  
PR 17-AUG-1999; 99US-0147303P.  
PR 18-AUG-1999; 99US-0147303P.  
PR 19-AUG-1999; 99US-0147303P.  
PR 20-AUG-1999; 99US-0147303P.  
PR 21-AUG-1999; 99US-0147303P.  
PR 22-AUG-1999; 99US-0147303P.  
PR 23-AUG-1999; 99US-0147303P.  
PR 24-AUG-1999; 99US-0147303P.  
PR 25-AUG-1999; 99US-0147303P.  
PR 26-AUG-1999; 99US-0147303P.  
PR 27-AUG-1999; 99US-0147303P.  
PR 28-AUG-1999; 99US-0147303P.  
PR 29-AUG-1999; 99US-0147303P.  
PR 30-AUG-1999; 99US-0147303P.  
PR 31-AUG-1999; 99US-0147303P.  
PR 01-SEP-1999; 99US-0147303P.  
PR 02-SEP-1999; 99US-0147303P.  
PR 03-SEP-1999; 99US-0147303P.  
PR 04-SEP-1999; 99US-0147303P.  
PR 05-SEP-1999; 99US-0147303P.  
PR 06-SEP-1999; 99US-0147303P.  
PR 07-SEP-1999; 99US-0147303P.  
PR 08-SEP-1999; 99US-0147303P.  
PR 09-SEP-1999; 99US-0147303P.  
PR 10-SEP-1999; 99US-0147303P.  
PR 11-SEP-1999; 99US-0147303P.  
PR 12-SEP-1999; 99US-0147303P.  
PR 13-SEP-1999; 99US-0147303P.  
PR 14-SEP-1999; 99US-0147303P.  
PR 15-SEP-1999; 99US-0147303P.  
PR 16-SEP-1999; 99US-0147303P.  
PR 17-SEP-1999; 99US-0147303P.  
PR 18-SEP-1999; 99US-0147303P.  
PR 19-SEP-1999; 99US-0147303P.  
PR 20-SEP-1999; 99US-0147303P.  
PR 21-SEP-1999; 99US-0147303P.  
PR 22-SEP-1999; 99US-0147303P.  
PR 23-SEP-1999; 99US-0147303P.  
PR 24-SEP-1999; 99US-0147303P.  
PR 25-SEP-1999; 99US-0147303P.  
PR 26-SEP-1999; 99US-0147303P.  
PR 27-SEP-1999; 99US-0147303P.  
PR 28-SEP-1999; 99US-0147303P.  
PR 29-SEP-1999; 99US-0147303P.  
PR 30-SEP-1999; 99US-0147303P.  
PR 01-OCT-1999; 99US-0147303P.  
PR 02-OCT-1999; 99US-0147303P.  
PR 03-OCT-1999; 99US-0147303P.  
PR 04-OCT-1999; 99US-0147303P.  
PR 05-OCT-1999; 99US-0147303P.  
PR 06-OCT-1999; 99US-0147303P.  
PR 07-OCT-1999; 99US-0147303P.  
PR 08-OCT-1999; 99US-0147303P.  
PR 09-OCT-1999; 99US-0147303P.  
PR 10-OCT-1999; 99US-0147303P.  
PR 11-OCT-1999; 99US-0147303P.  
PR 12-OCT-1999; 99US-0147303P.  
PR 13-OCT-1999; 99US-0147303P.  
PR 14-OCT-1999; 99US-0147303P.  
PR 15-OCT-1999; 99US-0147303P.  
PR 16-OCT-1999; 99US-0147303P.  
PR 17-OCT-1999; 99US-0147303P.  
PR 18-OCT-1999; 99US-0147303P.  
PR 19-OCT-1999; 99US-0147303P.  
PR 20-OCT-1999; 99US-0147303P.  
PR 21-OCT-1999; 99US-0147303P.  
PR 22-OCT-1999; 99US-0147303P.  
PR 23-OCT-1999; 99US-0147303P.  
PR 24-OCT-1999; 99US-0147303P.  
PR 25-OCT-1999; 99US-0147303P.  
PR 26-OCT-1999; 99US-0147303P.  
PR 27-OCT-1999; 99US-0147303P.  
PR 28-OCT-1999; 99US-0147303P.  
PR 29-OCT-1999; 99US-0147303P.  
PR 30-OCT-1999; 99US-0147303P.

Mon Sep 27 08:51:16 2004

```
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161922P.
PR 28-OCT-1999; 99US-0161993P.
PR 29-OCT-1999; 99US-0162142P.

Query Match      29.3%; Score 516.5; DB 3; Length 382;
Best Local Similarity 32.2%; Pred. NO. 1.9e-46;
Matches 115; Conservative 81; Mismatches 132; Indels 29; Gaps 7;

QY 1 MVK-----ALVIINCILAIAGNCGPLIMRLYFNNGKRIWFSTFETAGFEVFIPL 53
Db 16 MVKPPVKRSLTLIVTYFFLFGSIASSLLAKYIFVYGGSSRWSTWVQSAGPELLILI 75
QY 54 LFSY----LTRRRSNNVGDSTSFLLIKPRLIAAVIVGILSGFDNYLYAGIAYLPVSTA 109
Db 76 YFPHYVLKTTTR-----PFTRFTL--RHLIFSVLIGLVGFNNFLFSWGTSLPVSTS 127
QY 110 ALIIASQLAPIAFSPFMVKHKFTPTTINAVULLTVGAAVLGMHTETDKPVHETHKQYIT 169
Db 128 SLLLSQLVFTLILSRILVVKQITFSNLCVLLTSSVLLALDSSKDXPSGLTKXYFI 187
QY 170 GELITVAAAVMVAFILPLVELAYOKAKOTMSYTLVLEFQLILCLLASIVSVIGMPIAGDF 229
Db 188 GYVSTIGAGLLFALYLPVTEKLY---RTVVCYAMWVQLVMEFAATVFATIGMACEGGF 244
QY 230 KQALPKAREFEKLGALFYVVAVFSAIL-WQGFFLGALGLIFSTSLVSGIMISVLLPIT 288
Db 245 KEMVKEANHVTGKPTFYMTFAILANVTVQLSFAATSGMVTLSGTGGICMTALLAMN 304
QY 289 EVLAVIFYHEKFOAEKGLSLALSLWGVSVYFYGEIKSGEDKRRIOQESQETEQSSL 345
Db 305 VIGGWAYGDVFGVKIVSTVLCINGFSSYTYGMY-----MKMKKEEKEGEYSKV 355
```

Search completed: September 23, 2004, 23:47:15  
Job time : 91 secs

















Mon Sep 27 08:51:13 2004

us-09-913-767-1.oligo.rnpb

Sequence 3028, Ap  
Sequence 3029, Ap  
Sequence 8034, Ap  
Sequence 750, App  
Sequence 1943, Ap  
Sequence 3788, Ap  
Sequence 14076, A  
Sequence 13064, A  
Sequence 1651, Ap  
Sequence 140569, A  
Sequence 13684, A  
Sequence 166, App  
Sequence 64468, A  
Sequence 296420, A  
Sequence 296421, A  
Sequence 296422, A  
Sequence 64468, A  
Sequence 296420, A  
Sequence 296421, A  
Sequence 296422, A  
Sequence 758, App  
Sequence 13548, A  
Sequence 310965, A  
Sequence 310965, A  
Sequence 2769, Ap  
Sequence 8203, Ap  
Sequence 370, App  
Sequence 286, App  
Sequence 370, App  
Sequence 370, App  
Sequence 105095, A  
Sequence 105096, A  
Sequence 129799, A  
Sequence 105095, A  
Sequence 105096, A  
Sequence 129799, A  
Sequence 20143, A

16 1.3 494 16 US-10-027-632-3028  
16 1.3 494 16 US-10-027-632-3029  
16 1.3 495 9 US-09-783-590-8034  
16 1.3 495 12 US-10-152-319A-750  
16 1.3 495 13 US-10-282-122A-1943  
16 1.3 496 10 US-09-918-995-3788  
16 1.3 496 10 US-09-918-995-14076  
16 1.3 499 10 US-09-814-353-13064  
16 1.3 499 10 US-09-991-936-1651  
16 1.3 500 10 US-10-424-599-140569  
16 1.3 501 13 US-09-814-353-13684  
16 1.3 501 15 US-10-116-712-166  
16 1.3 502 13 US-10-027-632-64468  
16 1.3 502 13 US-10-027-632-296420  
16 1.3 502 13 US-10-027-632-296421  
16 1.3 502 13 US-10-027-632-296422  
16 1.3 502 13 US-10-027-632-64468  
16 1.3 502 16 US-10-027-632-296420  
16 1.3 502 16 US-10-027-632-296421  
16 1.3 502 16 US-10-027-632-296422  
16 1.3 502 9 US-09-924-035A-758  
16 1.3 503 13 US-10-424-599-13548  
16 1.3 503 13 US-10-027-632-310965  
16 1.3 503 16 US-10-027-632-310965  
16 1.3 503 17 US-10-767-795-2769  
16 1.3 505 9 US-09-864-761-8203  
16 1.3 506 9 US-09-920-300A-370  
16 1.3 506 10 US-09-770-961-286  
16 1.3 506 14 US-10-033-538-370  
16 1.3 506 15 US-10-099-946-370  
16 1.3 507 13 US-10-027-632-105095  
16 1.3 507 13 US-10-027-632-105096  
16 1.3 507 13 US-10-027-632-129799  
16 1.3 507 16 US-10-027-632-105095  
16 1.3 507 16 US-10-027-632-105096  
16 1.3 507 16 US-10-027-632-129799  
16 1.3 514 13 US-10-425-114-20143

ALIGNMENTS

RESULT 1  
US-10-085-783A-54761  
; Sequence 54761, Application US/10085783A  
; Publication No. US20040037841A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2002  
; CURRENT APPLICATION NUMBER: US/10/085,783A  
; CURRENT FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 54761  
; LENGTH: 470  
; TYPE: DNA  
; ORGANISM: Human  
US-10-085-783A-54761

Query Match 1.7%; Score 21; DB 13; Length 470;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1164 TTGTGTTTGTATTGTT 1184  
|||||

Db 8 TTGTGTTTGTATTGTT 28  
RESULT 2  
US-10-242-535A-54761  
; Sequence 54761, Application US/10242535A  
; Publication No. US20040013663A1  
; GENERAL INFORMATION:  
; APPLICANT: ChondroGene Inc.  
; APPLICANT: Liew, C.C.  
; TITLE OF INVENTION: Compositions and Methods Relating to Osteoarthritis  
; FILE REFERENCE: 4231/2005  
; CURRENT APPLICATION NUMBER: US/10/242,535A  
; CURRENT FILING DATE: 2002-09-12  
; PRIOR APPLICATION NUMBER: US 10/085,783  
; PRIOR FILING DATE: 2002-02-28  
; PRIOR APPLICATION NUMBER: US 60/305,340  
; PRIOR FILING DATE: 2001-07-13  
; PRIOR APPLICATION NUMBER: US 60/275,017  
; PRIOR FILING DATE: 2001-03-12  
; PRIOR APPLICATION NUMBER: US 60/271,955  
; PRIOR FILING DATE: 2001-02-28  
; NUMBER OF SEQ ID NOS: 58994  
; SOFTWARE: PatentIn version 3.2  
; SEQ ID NO 54761  
; LENGTH: 470  
; TYPE: DNA  
; ORGANISM: Human  
US-10-242-535A-54761

Query Match 1.7%; Score 21; DB 16; Length 470;  
Best Local Similarity 100.0%; Pred. No. 5.7;  
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;  
QY 1164 TTGTGTTTGTATTGTT 1184  
|||||

Db 8 TTGTGTTTGTATTGTT 28

RESULT 3  
US-10-282-122A-31258  
; Sequence 31258, Application US/10282122A  
; Publication No. US20040029129A1  
; GENERAL INFORMATION:  
; APPLICANT: Wang, Liangsu  
; APPLICANT: Zamudio, Carlos  
; APPLICANT: Malone, Cheryl  
; APPLICANT: Haselbeck, Robert  
; APPLICANT: Ohlsen, Kari  
; APPLICANT: Zyskind, Judith  
; APPLICANT: Wall, Daniel  
; APPLICANT: Trawick, John  
; APPLICANT: Carr, Grant  
; APPLICANT: Yamamoto, Robert  
; APPLICANT: Forsyth, R.  
; APPLICANT: Xu, H.  
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms  
; FILE REFERENCE: ELIIRA.034A  
; CURRENT APPLICATION NUMBER: US/10/282,122A  
; CURRENT FILING DATE: 2003-02-20  
; PRIOR APPLICATION NUMBER: 60/191,078  
; PRIOR FILING DATE: 2000-03-21  
; PRIOR APPLICATION NUMBER: 60/206,848  
; PRIOR FILING DATE: 2000-05-23  
; PRIOR APPLICATION NUMBER: 60/207,727  
; PRIOR FILING DATE: 2000-05-26  
; PRIOR APPLICATION NUMBER: 60/230,335  
; PRIOR FILING DATE: 2000-09-06  
; PRIOR APPLICATION NUMBER: 60/230,347  
; PRIOR FILING DATE: 2000-09-09  
; PRIOR APPLICATION NUMBER: 60/242,578  
; PRIOR FILING DATE: 2000-10-23  
; PRIOR APPLICATION NUMBER: 60/253,625

```
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 31258
; LENGTH: 729
; TYPE: DNA
; ORGANISM: Pasteurella multocida
US-10-282-122A-31258

Query Match      1.7%; Score 21; DB 13; Length 729;
Best Local Similarity 100.0%; Pred. No. 5.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AAGATGAAGAATGTTTGATA 42
Db 290 AAGATGAAGAATGTTTGATA 310

RESULT 4
US-10-425-114-23082
; Sequence 23082, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated with
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 23082
; LENGTH: 1093
; TYPE: DNA
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB3594-005-F3_FLI
US-10-425-114-23082

Query Match      1.7%; Score 21; DB 13; Length 1093;
Best Local Similarity 100.0%; Pred. No. 5.9;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1164 TTGTGTGTTTTTCTTATTGTT 1184
Db 993 TTGTGTGTTTTTCTTATTGTT 1013

RESULT 5
US-10-398-221-3618/c
; Sequence 3618, Application US/10398221
; Publication No. US20040018514A1
; GENERAL INFORMATION:
; APPLICANT: KUNST, Frederik
; APPLICANT: GLASER, Philippe
; TITLE OF INVENTION: Listeria innocua, genome and applications
; FILE REFERENCE: 344 702 - US
; CURRENT APPLICATION NUMBER: US/10/398,221
; CURRENT FILING DATE: 2003-03-27
; PRIOR APPLICATION NUMBER: PCT/FR 01/03 061
; PRIOR FILING DATE: 2001-10-04
; PRIOR APPLICATION NUMBER: FR 00/12 697
```

```
; PRIOR FILING DATE: 2000-10-04
; NUMBER OF SEQ ID NOS: 4025
; SOFTWARE: PatentIn version 3.0
; SEQ ID NO 3618
; LENGTH: 2791
; TYPE: DNA
; ORGANISM: Listeria monocytogenes 4b
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(end)
; OTHER INFORMATION: n can be any nucleotide: a,g,c or t/u
US-10-398-221-3618

Query Match      1.7%; Score 21; DB 16; Length 2791;
Best Local Similarity 100.0%; Pred. No. 6.2;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 685 GCTACTTTTCTGTGTCATT 705
Db 693 GCTACTTTTCTGTGTCATT 673

RESULT 6
US-10-087-192-2053
; Sequence 2053, Application US/10087192
; Publication No. US20020182586A1
; GENERAL INFORMATION:
; APPLICANT: Morris, David W.
; APPLICANT: Engelhard, Eric K.
; TITLE OF INVENTION: NOVEL COMPOSITIONS AND METHODS FOR
; TITLE OF INVENTION: CANCER
; FILE REFERENCE: 529452000122
; CURRENT APPLICATION NUMBER: US/10/087,192
; CURRENT FILING DATE: 2002-03-01
; PRIOR APPLICATION NUMBER: US 09/747,377
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: US 09/798,586
; PRIOR FILING DATE: 2001-03-02
; NUMBER OF SEQ ID NOS: 2059
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 2053
; LENGTH: 23126
; TYPE: DNA
; ORGANISM: Mus musculus
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (1)...(23126)
; OTHER INFORMATION: n = A,T,C or G
US-10-087-192-2053

Query Match      1.7%; Score 21; DB 13; Length 23126;
Best Local Similarity 100.0%; Pred. No. 6.8;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 405 TCTCTTCGCTTCTTCTGTAGT 425
Db 5281 TCTCTTCGCTTCTTCTGTAGT 5301

RESULT 7
US-10-322-281-820
; Sequence 820, Application US/10322281
; Publication No. US20040126762A1
; GENERAL INFORMATION:
; APPLICANT: David W. Morris
; APPLICANT: Marc S. Malandro
; TITLE OF INVENTION: Novel Compositions and Methods in Cancer
; FILE REFERENCE: 529452001000
; CURRENT APPLICATION NUMBER: US/10/322,281
; CURRENT FILING DATE: 2002-12-17
; NUMBER OF SEQ ID NOS: 866
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 820
```

Mon Sep 27 08:51:13 2004

us-09-913-767-1.oligo.rnpb

```

; LENGTH: 41787
; TYPE: DNA
; ORGANISM: Homo sapiens
; US-10-322-281-820

Query Match
; 1.7%; Score 21; DB 17; Length 41787;
; Best Local Similarity 100.0%; Pred. No. 7;
; Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1164 TTGCTGTTTGGTTTATTTGTT 1184
Db 31795 TTGCTGTTTGGTTTATTTGTT 31815

RESULT 8
US-10-027-632-121522
; Sequence 121522, Application US/10027632
; Publication No. US20020198371A1
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; PRIOR APPLICATION NUMBER: US 60/185,218
; PRIOR FILING DATE: 2000-02-24
; PRIOR APPLICATION NUMBER: US 60/167,363
; PRIOR FILING DATE: 1999-11-23
; PRIOR APPLICATION NUMBER: US 60/156,358
; PRIOR FILING DATE: 1999-09-28
; PRIOR APPLICATION NUMBER: US 60/146,002
; PRIOR FILING DATE: 1999-08-09
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121522
; LENGTH: 854
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-121522

Query Match
; 1.6%; Score 20; DB 13; Length 854;
; Best Local Similarity 100.0%; Pred. No. 20;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 AAGTATGATATAAAATTGAA 1213
Db 43 AAGTATGATATAAAATTGAA 62

RESULT 9
US-10-027-632-121522
; Sequence 121522, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121522
; LENGTH: 854
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-121522

Query Match
; 1.6%; Score 20; DB 13; Length 854;
; Best Local Similarity 100.0%; Pred. No. 20;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 AAGTATGATATAAAATTGAA 1213
Db 43 AAGTATGATATAAAATTGAA 62

RESULT 9
US-10-027-632-121522
; Sequence 121522, Application US/10027632
; Publication No. US20030204075A9
; GENERAL INFORMATION:
; APPLICANT: Wang, David G.
; TITLE OF INVENTION: Identification and Mapping of Single Nucleotide
; Polymorphisms in the Human Genome
; FILE REFERENCE: 108827.129
; CURRENT APPLICATION NUMBER: US/10/027,632
; CURRENT FILING DATE: 2002-04-30
; PRIOR APPLICATION NUMBER: US 60/218,006
; PRIOR FILING DATE: 2000-07-12
; PRIOR APPLICATION NUMBER: US 60/198,676
; PRIOR FILING DATE: 2000-04-20
; PRIOR APPLICATION NUMBER: US 60/193,483
; PRIOR FILING DATE: 2000-03-29
; NUMBER OF SEQ ID NOS: 325720
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 121522
; LENGTH: 854
; TYPE: DNA
; ORGANISM: Human
; US-10-027-632-121522

Query Match
; 1.6%; Score 20; DB 16; Length 854;
; Best Local Similarity 100.0%; Pred. No. 20;
; Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1194 AAGTATGATATAAAATTGAA 1213
Db 43 AAGTATGATATAAAATTGAA 62

RESULT 10
US-10-282-122A-24401/C
; Sequence 24401, Application US/10282122A
; Publication No. US20040029129A1
; GENERAL INFORMATION:
; APPLICANT: Wang, Liangsu
; APPLICANT: Zamudio, Carlos
; APPLICANT: Malone, Cheryl
; APPLICANT: Haselbeck, Robert
; APPLICANT: Ohlsen, Kari
; APPLICANT: Zyskind, Judith
; APPLICANT: Wall, Daniel
; APPLICANT: Trawick, John
; APPLICANT: Carr, Grant
; APPLICANT: Yamamoto, Robert
; APPLICANT: Forsyth, R.
; APPLICANT: Xu, H.
; TITLE OF INVENTION: Identification of Essential Genes in Microorganisms
; FILE REFERENCE: ELITRA.034A
; CURRENT APPLICATION NUMBER: US/10/282,122A
; CURRENT FILING DATE: 2003-02-20
; PRIOR APPLICATION NUMBER: 60/191,078
; PRIOR FILING DATE: 2000-03-21
; PRIOR APPLICATION NUMBER: 60/206,848
; PRIOR FILING DATE: 2000-05-23
; PRIOR APPLICATION NUMBER: 60/207,727
; PRIOR FILING DATE: 2000-05-26
; PRIOR APPLICATION NUMBER: 60/230,335
; PRIOR FILING DATE: 2000-09-06
; PRIOR APPLICATION NUMBER: 60/230,347
; PRIOR FILING DATE: 2000-09-09
; PRIOR APPLICATION NUMBER: 60/242,578
; PRIOR FILING DATE: 2000-10-23
; PRIOR APPLICATION NUMBER: 60/253,625
; PRIOR FILING DATE: 2000-11-27
; PRIOR APPLICATION NUMBER: 60/257,931
; PRIOR FILING DATE: 2000-12-22
; PRIOR APPLICATION NUMBER: 60/267,636
; PRIOR FILING DATE: 2001-02-09
; PRIOR APPLICATION NUMBER: 60/269,308
; PRIOR FILING DATE: 2001-02-16
; Remaining Prior Application data removed - See File Wrapper or PALM.
; NUMBER OF SEQ ID NOS: 78614
; SOFTWARE: PatentIn version 3.1
; SEQ ID NO 24401
; LENGTH: 1017
; TYPE: DNA
; ORGANISM: Listeria monocytogenes

```

US-10-282-122A-24401

Query Match 1.6%; Score 20; DB 13; Length 1017;  
Best Local Similarity 100.0%; Pred. No. 20;  
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 685 GCTACTTTTCTGTGTCAT 704  
|||||  
DB 20 GCTACTTTTCTGTGTCAT 1

RESULT 11

US-10-424-599-42537  
; Sequence 42537, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 42537  
; LENGTH: 317  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_138411C.1  
US-10-424-599-42537

Query Match 1.6%; Score 19; DB 13; Length 317;  
Best Local Similarity 100.0%; Pred. No. 62;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 13 AGCAAGAAGAAGATGAAGA 31  
|||||  
DB 56 AGCAAGAAGAAGATGAAGA 74

RESULT 12

US-10-424-599-123715/c  
; Sequence 123715, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 123715  
; LENGTH: 411  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_82722C.1  
US-10-424-599-123715

Query Match 1.6%; Score 19; DB 13; Length 411;  
Best Local Similarity 100.0%; Pred. No. 63;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 AGAAGAAGATGAAGATGG 35  
|||||  
DB 28 AGAAGAAGATGAAGATGG 10

RESULT 13

US-10-425-114-9121/c  
; Sequence 9121, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E.  
; APPLICANT: Tabaska, Jack E.  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53113)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 9121  
; LENGTH: 802  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700834123\_FLI  
US-10-425-114-9121

Query Match 1.6%; Score 19; DB 13; Length 802;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 AGAAGAAGATGAAGATGG 35  
|||||  
DB 37 AGAAGAAGATGAAGATGG 19

RESULT 14

US-10-424-599-4662  
; Sequence 4662, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 4662  
; LENGTH: 935  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_104214C.1  
US-10-424-599-4662

Query Match 1.6%; Score 19; DB 13; Length 935;  
Best Local Similarity 100.0%; Pred. No. 65;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 17 AGAAGAAGATGAAGATGG 35  
|||||  
DB 899 AGAAGAAGATGAAGATGG 917

RESULT 15

US-09-887-576-749  
; Sequence 749, Application US/09887576  
; Patent No. US20020144047A1  
; GENERAL INFORMATION:

Mon Sep 27 08:51:13 2004

```

; APPLICANT: Budworth, P.
; APPLICANT: Brown, D.
; APPLICANT: Chang, H.
; APPLICANT: Zhu, T.
; APPLICANT: Han, B.
; APPLICANT: Wang, X.
; APPLICANT: Cooper, X.
; TITLE OF INVENTION: Promoters for regulation of plant expression
; FILE REFERENCE: 1360.001US1
; CURRENT APPLICATION NUMBER: US/09/887,576
; CURRENT FILING DATE: 2001-06-25
; PRIOR APPLICATION NUMBER: US 60/213,848
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/214,087
; PRIOR FILING DATE: 2000-06-23
; PRIOR APPLICATION NUMBER: US 60/258,692
; PRIOR FILING DATE: 2000-12-29
; NUMBER OF SEQ ID NOS: 875
; SOFTWARE: FastSeq for Windows Version 4.0
; SEQ ID NO 749
; LENGTH: 1200
; TYPE: DNA
; ORGANISM: Arabidopsis thaliana
US-09-887-576-749

```

```

Query Match      1.6%; Score 19; DB 9; Length 1200;
Best Local Similarity 100.0%; Pred. No. 66;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```

```

QY      785 ATGCATTGATGATGATCAC 803
          |||||
DB      282 ATGCATTGATGATGATCAC 300

```

Search completed: September 25, 2004, 15:55:53  
Job time : 681 secs



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2004, 11:28:06 ; Search time 107 Seconds  
(without alignments)  
6353.410 Million cell updates/sec

Title: US-09-913-767-1  
Perfect score: 1225  
Sequence: 1 aaacagcagcagcagaa.....aattgaaagattgagct 1225

Scoring table: OLIGO NUC

Gapop\_60.0 , Gapext 60.0

Searched: 682709 seqs, 277475446 residues

Word size : 0

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 200000000

Post-processing: Listing first 1000 summaries

Database : Issued Patents NA:\*

- 1: /cgn2\_6/prodata/2/ina/5A COMB.seq.\*
- 2: /cgn2\_6/prodata/2/ina/5B COMB.seq.\*
- 3: /cgn2\_6/prodata/2/ina/6A COMB.seq.\*
- 4: /cgn2\_6/prodata/2/ina/6B COMB.seq.\*
- 5: /cgn2\_6/prodata/2/ina/PCFUS COMB.seq.\*
- 6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
C 1	19	1.6	500	US-09-621-976-15356	Sequence 15356, A
C 2	19	1.6	3647	US-08-588-983-13	Sequence 13, Appl
C 3	19	1.6	3647	US-08-588-976-13	Sequence 13, Appl
C 4	18	1.5	304	US-09-712-016-15	Sequence 15, Appl
C 5	18	1.5	412	US-09-621-976-15361	Sequence 15361, A
C 6	18	1.5	413	US-09-621-976-15359	Sequence 15359, A
C 7	18	1.5	413	US-09-621-976-15367	Sequence 15367, A
C 8	18	1.5	414	US-09-621-976-15364	Sequence 15364, A
C 9	18	1.5	415	US-09-621-976-15358	Sequence 15358, A
C 10	18	1.5	416	US-09-621-976-15360	Sequence 15360, A
C 11	18	1.5	416	US-09-621-976-15362	Sequence 15362, A
C 12	18	1.5	425	US-09-621-976-15366	Sequence 15366, A
C 13	18	1.5	434	US-09-621-976-15365	Sequence 15365, A
C 14	18	1.5	914	US-09-227-357-123	Sequence 123, Appl
C 15	18	1.5	1236	US-08-741-134-5	Sequence 5, Appl
C 16	18	1.5	1433	US-09-252-991A-16282	Sequence 16282, A
C 17	18	1.5	1833	US-09-252-991A-15995	Sequence 15995, A
C 18	18	1.5	2532	US-09-252-991A-16097	Sequence 16097, A
C 19	18	1.5	4460	US-09-103-875-4	Sequence 4, Appl
C 20	18	1.5	5562	US-10-204-708-63	Sequence 63, Appl
C 21	18	1.5	8532	US-08-452-655B-1	Sequence 1, Appl
C 22	18	1.5	8532	US-08-450-582-1	Sequence 1, Appl
C 23	18	1.5	9606	US-07-741-940-1	Sequence 1, Appl
C 24	18	1.5	9606	US-08-289-548A-1	Sequence 1, Appl
C 25	18	1.5	9606	US-08-452-654-1	Sequence 1, Appl
C 26	18	1.5	9606	US-08-370-235A-1	Sequence 1, Appl
C 27	18	1.5	9606	US-08-449-731-1	Sequence 1, Appl

18	1.5	40000	4	US-09-780-049-18	Sequence 18, Appl
18	1.5	640681	4	US-09-790-988-1	Sequence 1, Appl
C 29	17	1.4	459	US-09-621-976-2386	Sequence 2386, Ap
C 30	17	1.4	570	US-09-543-681A-3616	Sequence 3616, Ap
C 31	17	1.4	700	US-09-236-097-11	Sequence 11, Appl
C 32	17	1.4	723	US-09-328-475C-323	Sequence 323, App
C 33	17	1.4	744	US-09-328-475C-322	Sequence 322, App
C 34	17	1.4	810	US-09-107-532A-2028	Sequence 2028, Ap
C 35	17	1.4	817	US-09-533-029-43	Sequence 43, Appl
C 36	17	1.4	822	US-09-328-475C-187	Sequence 187, App
C 37	17	1.4	873	US-09-134-000C-311	Sequence 311, App
C 38	17	1.4	965	US-09-141-135-1	Sequence 1, Appl
C 39	17	1.4	965	US-09-533-029-97	Sequence 97, Appl
C 40	17	1.4	1024	US-09-328-475C-3	Sequence 3, Appl
C 41	17	1.4	1032	US-09-252-991A-9669	Sequence 9669, Ap
C 42	17	1.4	1233	US-09-252-991A-9747	Sequence 9747, Ap
C 43	17	1.4	1248	US-09-328-352-573	Sequence 573, App
C 44	17	1.4	1425	US-08-883-515-1	GENERAL INFORMAT
C 45	17	1.4	1458	US-09-108-010B-12	Sequence 12, Appl
C 46	17	1.4	1516	US-09-976-594-599	Sequence 599, App
C 47	17	1.4	1539	US-09-461-325-50	Sequence 50, Appl
C 48	17	1.4	1539	US-10-012-542-50	Sequence 50, Appl
C 49	17	1.4	1681	US-09-461-325-115	Sequence 115, App
C 50	17	1.4	1681	US-10-012-542-115	Sequence 115, App
C 51	17	1.4	1895	US-09-620-312D-397	Sequence 397, App
C 52	17	1.4	1954	US-08-097-829-3	Sequence 3, Appl
C 53	17	1.4	1954	US-08-577-403-3	Sequence 3, Appl
C 54	17	1.4	1991	US-08-295-814B-9	Sequence 9, Appl
C 55	17	1.4	1991	US-09-343-361-9	Sequence 9, Appl
C 56	17	1.4	1991	PCT-US93-01959-9	Sequence 9, Appl
C 57	17	1.4	2262	US-08-674-887A-5	Sequence 5, Appl
C 58	17	1.4	2262	US-08-951-844-5	Sequence 5, Appl
C 59	17	1.4	2262	US-09-167-109-2	Sequence 2, Appl
C 60	17	1.4	2262	US-09-412-347-5	Sequence 5, Appl
C 61	17	1.4	2420	US-08-890-094-3	Sequence 3, Appl
C 62	17	1.4	3153	US-09-614-474-1	Sequence 1, Appl
C 63	17	1.4	3159	US-09-566-921-16	Sequence 16, Appl
C 64	17	1.4	3184	US-09-614-474-12	Sequence 12, Appl
C 65	17	1.4	3339	US-09-566-921-15	Sequence 15, Appl
C 66	17	1.4	3357	US-08-726-214-7	Sequence 7, Appl
C 67	17	1.4	3518	US-09-412-210-2	Sequence 2, Appl
C 68	17	1.4	3518	US-10-121-911A-2	Sequence 2, Appl
C 69	17	1.4	3911	US-09-423-890-1	Sequence 1, Appl
C 70	17	1.4	4693	US-09-359-756-1	Sequence 1, Appl
C 71	17	1.4	4925	US-08-470-260-4	Sequence 4, Appl
C 72	17	1.4	5925	US-08-471-491-4	Sequence 4, Appl
C 73	17	1.4	5925	US-08-466-662-4	Sequence 4, Appl
C 74	17	1.4	6702	US-08-987-439-2	Sequence 2, Appl
C 75	17	1.4	7827	US-09-620-312D-104	Sequence 104, App
C 76	17	1.4	10299	US-08-477-451-1	Sequence 1, Appl
C 77	17	1.4	10299	US-08-477-451-5	Sequence 5, Appl
C 78	17	1.4	11014	US-08-956-171B-91	Sequence 91, Appl
C 79	17	1.4	11485	US-09-410-464-9	Sequence 9, Appl
C 80	17	1.4	15462	US-09-073-492-1	Sequence 1, Appl
C 81	17	1.4	19932	US-08-477-451-25	Sequence 25, Appl
C 82	17	1.4	392000	US-10-027-983-11	Sequence 11, Appl
C 83	17	1.4	1230025	US-09-198-452A-1	Sequence 1, Appl
C 84	17	1.4	1664976	US-08-916-421B-1	Sequence 1, Appl
C 85	17	1.3	20	US-09-198-452A-6495	Sequence 6495, Ap
C 86	16	1.3	255	US-09-107-532A-1600	Sequence 1600, Ap
C 87	16	1.3	256	US-09-023-655-474	Sequence 474, App
C 88	16	1.3	256	US-09-345-882-21	Sequence 21, Appl
C 89	16	1.3	258	US-09-172-108-52	Sequence 52, Appl
C 90	16	1.3	293	US-08-172-711-54	Sequence 54, Appl
C 91	16	1.3	293	US-09-443-184-14	Sequence 14, Appl
C 92	16	1.3	309	US-09-702-705-1222	Sequence 1222, Ap
C 93	16	1.3	309	US-09-736-457-1222	Sequence 1222, Ap
C 94	16	1.3	309	US-09-614-124B-1222	Sequence 1222, Ap
C 95	16	1.3	309	US-09-671-325-1222	Sequence 1222, Ap
C 96	16	1.3	331	US-09-624-268B-10	Sequence 10, Appl
C 97	16	1.3	332	US-09-624-268B-11	Sequence 11, Appl
C 98	16	1.3	446	US-08-256-077-5	Sequence 5, Appl
C 99	16	1.3	446	US-08-466-127-5	Sequence 5, Appl
C 100	16	1.3	446		















```

RESULT 2
US-08-588-983-13/c
; Sequence 13, Application US/08588983
; Patent No. 5854067
; GENERAL INFORMATION:
; APPLICANT: Christopher B. Newgard, et al.
; TITLE OF INVENTION: Methods and Compositions
; TITLE OF INVENTION: for Inhibiting Hexokinase
; NUMBER OF SEQUENCES: 43
; CORRESPONDENCE ADDRESS:
; ADDRESSEE: Arnold, White & Durkee
; STREET: P. O. Box 4433
; CITY: Houston
; STATE: TX
; COUNTRY: US
;

```

CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Fussey, Shelley P.M.  
REGISTRATION NUMBER: 39,458  
REFERENCE/DOCKET NUMBER: UTSD:481/FUS  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (512) 418-3000  
TELEFAX: (512) 474-7577  
TELEX: n/a  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 3647 base pairs  
TYPE: nucleic acid  
STRANDEDNESS: single



```
; TOPOLOGY: linear
US-08-588-976-13

Query Match          1.6%; Score 19; DB 2; Length 3647;
Best Local Similarity 100.0%; Pred. No. 9.2;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 184 TTGTCCTCCTTCTCAGCC 202
Db 1602 TTGTCCTCCTTCTCAGCC 1584

RESULT 4
US-09-712-016-15/c
; Sequence 15, Application US/09712016
; Patent No. 6468790
; GENERAL INFORMATION:
; APPLICANT: Giese, Klaus
; TITLE OF INVENTION: METASTATIC BREAST AND COLON CANCER
; FILE REFERENCE: 200130.460
; CURRENT APPLICATION NUMBER: US/09/712.016
; PRIOR FILING DATE: 2000-11-13
; PRIOR FILING DATE: PRIOR FILING DATE: 1999-10-13
; NUMBER OF SEQ ID NOS: 85
; SOFTWARE: FastSeq for Windows Version 3.0
; SEQ ID NO 15
; LENGTH: 304
; TYPE: DNA
; ORGANISM: Homo sapiens
; NAME/KEY: misc feature
; LOCATION: (1)...(304)
; OTHER INFORMATION: n = A, T, C or G
US-09-712-016-15

Query Match          1.5%; Score 18; DB 4; Length 304;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 40 ATAATCATAACTGATT 57
Db 147 ATAATCATAACTGATT 130

RESULT 5
US-09-621-976-15361/c
; Sequence 15361, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S. J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15361
; LENGTH: 412
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15361

Query Match          1.5%; Score 18; DB 4; Length 412;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 AGAAGCAAGAGAGTTCAA 761
Db 354 AGAAGCAAGAGAGTTCAA 337

RESULT 6
US-09-621-976-15359/c
; Sequence 15359, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15359
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15359

Query Match          1.5%; Score 18; DB 4; Length 413;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 AGAAGCAAGAGAGTTCAA 761
Db 354 AGAAGCAAGAGAGTTCAA 337

RESULT 7
US-09-621-976-15367/c
; Sequence 15367, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15367
; LENGTH: 413
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15367

Query Match          1.5%; Score 18; DB 4; Length 413;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 AGAAGCAAGAGAGTTCAA 761
Db 354 AGAAGCAAGAGAGTTCAA 337

RESULT 8
US-09-621-976-15364/c
; Sequence 15364, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621.976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
```

us-09-913-767-1.oligo.rni

Mon Sep 27 08:51:13 2004

```

; SOFTWARE: Patent.pm
; SEQ ID NO 15364
; LENGTH: 414
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15364

Query Match      1.5%; Score 18; DB 4; Length 414;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      744 AGAAGCAAGAGAGTTCAA 761
Db      354 AGAAGCAAGAGAGTTCAA 337

RESULT 9
US-09-621-976-15358/c
; Sequence 15358, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15358
; LENGTH: 415
; TYPE: DNA
; ORGANISM: Homo sapiens
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: 287_
; OTHER INFORMATION: n=a, g, c or t
US-09-621-976-15358

Query Match      1.5%; Score 18; DB 4; Length 415;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      744 AGAAGCAAGAGAGTTCAA 761
Db      355 AGAAGCAAGAGAGTTCAA 338

RESULT 10
US-09-621-976-15360/c
; Sequence 15360, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15360
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15360

Query Match      1.5%; Score 18; DB 4; Length 416;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      744 AGAAGCAAGAGAGTTCAA 761
Db      355 AGAAGCAAGAGAGTTCAA 338

RESULT 11
US-09-621-976-15362/c
; Sequence 15362, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15362
; LENGTH: 416
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15362

Query Match      1.5%; Score 18; DB 4; Length 416;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      744 AGAAGCAAGAGAGTTCAA 761
Db      354 AGAAGCAAGAGAGTTCAA 337

RESULT 12
US-09-621-976-15366/c
; Sequence 15366, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2
; CURRENT APPLICATION NUMBER: US/09/621,976
; CURRENT FILING DATE: 2000-07-21
; NUMBER OF SEQ ID NOS: 19335
; SOFTWARE: Patent.pm
; SEQ ID NO 15366
; LENGTH: 425
; TYPE: DNA
; ORGANISM: Homo sapiens
US-09-621-976-15366

Query Match      1.5%; Score 18; DB 4; Length 425;
Best Local Similarity 100.0%; Pred. No. 28;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY      744 AGAAGCAAGAGAGTTCAA 761
Db      355 AGAAGCAAGAGAGTTCAA 338

RESULT 13
US-09-621-976-15365/c
; Sequence 15365, Application US/09621976
; Patent No. 6639063
; GENERAL INFORMATION:
; APPLICANT: Dumas Milne Edwards, J.B.
; APPLICANT: Jobert, S.
; APPLICANT: Giordano, J.Y.
; TITLE OF INVENTION: ESTs and Encoded Human Proteins.
; FILE REFERENCE: GENSET.054PR2

```

; CURRENT APPLICATION NUMBER: US/09/621,976  
; CURRENT FILING DATE: 2000-07-21  
; NUMBER OF SEQ ID NOS: 19335  
; SOFTWARE: Patent.pm  
; SEQ ID NO 15365  
; LENGTH: 454  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
US-09-621-976-15365

Query Match 1.5% Score 18; DB 4; Length 454;  
Best Local Similarity 100.0%; Pred. No. 28;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 744 AGAAGCAAGAGAGTTCAA 761  
|||||  
DB 354 AGAAGCAAGAGAGTTCAA 337

RESULT 14  
US-09-227-357-123  
; Sequence 123, Application US/09227357  
; Patent No. 6342581  
; GENERAL INFORMATION:  
; APPLICANT: Fischer et al.  
; TITLE OF INVENTION: 123 Human Secreted Proteins  
; FILE REFERENCE: P2010P1  
; CURRENT APPLICATION NUMBER: US/09/227,357  
; CURRENT FILING DATE: 1999-01-08  
; EARLIER APPLICATION NUMBER: PCT/US98/13684  
; EARLIER FILING DATE: 1998-07-07  
; EARLIER APPLICATION NUMBER: 60/051,926  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,793  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,925  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,929  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,803  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,732  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,931  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,932  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,916  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,930  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,918  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,920  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,733  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/052,795  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,919  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/051,928  
; EARLIER FILING DATE: 1997-07-08  
; EARLIER APPLICATION NUMBER: 60/055,722  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,723  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,948  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,949  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,953  
; EARLIER FILING DATE: 1997-08-18

; EARLIER APPLICATION NUMBER: 60/055,950  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,947  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,964  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/056,360  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,684  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,984  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/055,954  
; EARLIER FILING DATE: 1997-08-18  
; EARLIER APPLICATION NUMBER: 60/058,785  
; EARLIER FILING DATE: 1997-09-12  
; EARLIER APPLICATION NUMBER: 60/058,664  
; EARLIER FILING DATE: 1997-09-12  
; EARLIER APPLICATION NUMBER: 60/058,660  
; EARLIER FILING DATE: 1997-09-12  
; EARLIER APPLICATION NUMBER: 60/058,661  
; EARLIER FILING DATE: 1997-09-12  
; NUMBER OF SEQ ID NOS: 672  
; SOFTWARE: Patentin Ver. 2.0  
; SEQ ID NO 123  
; LENGTH: 914  
; TYPE: DNA  
; ORGANISM: Homo sapiens  
; FEATURE:  
; NAME/KEY: SITE  
; LOCATION: (909)  
; OTHER INFORMATION: n equals a.t.g. or c  
US-09-227-357-123

Query Match 1.5% Score 18; DB 4; Length 914;  
Best Local Similarity 100.0%; Pred. No. 29;  
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 19 AAGAAGATGAAGATGCT 36  
|||||  
DB 576 AAGAAGATGAAGATGCT 593

RESULT 15  
US-08-741-134-5  
; Sequence 5, Application US/08741134  
; Patent No. 5861498  
; GENERAL INFORMATION:  
; APPLICANT: Litwack, Gerald  
; APPLICANT: Alnemri, Emad S.  
; APPLICANT: Fernandes-Alnemri, Teresa  
; TITLE OF INVENTION: IMMUNOPHILIN FKBP46 AND COMPOSITIONS FOR MAKING  
; TITLE OF INVENTION: AND  
; TITLE OF INVENTION: METHODS OF USING THE SAME  
; NUMBER OF SEQUENCES: 6  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Woodcock Washburn Kurtz Mackiewicz & No. 5861498ris  
; STREET: One Liberty Place - 46th floor  
; CITY: Philadelphia  
; STATE: Pennsylvania  
; COUNTRY: USA  
; ZIP: 19103  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: Windows 3.11  
; SOFTWARE: WordPerfect for Windows 6.1  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/741,134  
; FILING DATE:  
; CLASSIFICATION: 424  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/007,163

us-09-913-767-1.oligo.rni

Mon Sep 27 08:51:13 2004

```

; FILING DATE: 01-NOV-1995
; CLASSIFICATION: 424
; ATTORNEY/AGENT INFORMATION:
; NAME: Deluca, Mark
; REGISTRATION NUMBER: 33,229
; REFERENCE/DOCKET NUMBER: TJU-2090
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: 215-568-3100
; TELEFAX: 215-568-3439
; INFORMATION FOR SEQ ID NO: 5:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 1236 base pairs
; TYPE: nucleic acid
; STRANDEDNESS: double
; TOPOLOGY: not relevant
; MOLECULE TYPE: cDNA
; FEATURE:
; NAME/KEY: CDS
; LOCATION: 1..1236
; US-08-741-134-5

Query Match 1.5%; Score 18; DB 2; Length 1236;
Best Local Similarity 100.0%; Pred. No. 29;
Matches 18; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 15 CAAGAAGAGATGAGAA 32
Db 658 CAAGAAGAGATGAGAA 675

Search completed: September 25, 2004, 13:15:29
Job time : 127 secs

```

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2004, 08:27:06 ; Search time 106 Seconds  
(without alignments)  
6413.348 Million cell updates/sec

Title: US-09-913-767-1  
Perfect score: 1225  
Sequence: 1 aaacagcagcagcagcaaa.....aatttgaaagattattgagct 1225

Scoring table: IDENTITY NUC

Gapop 10.0 , Gapext 1.0

Searched: 682709 seqs, 277475446 residues

Total number of hits satisfying chosen parameters: 1365418

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents NA:\*  
1: /cgn2\_6/prodata/2/ina/5A-COMB.seq.\*  
2: /cgn2\_6/prodata/2/ina/5B-COMB.seq.\*  
3: /cgn2\_6/prodata/2/ina/6A-COMB.seq.\*  
4: /cgn2\_6/prodata/2/ina/6B-COMB.seq.\*  
5: /cgn2\_6/prodata/2/ina/PCUTS-COMB.seq.\*  
6: /cgn2\_6/prodata/2/ina/backfiles1.seq.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	48.6	4.0	7218	1	US-08-232-463-14
2	43	3.5	19233	4	US-10-204-708-45
3	40	3.3	6040	4	US-10-204-708-70
4	39.8	3.2	5360	4	US-10-204-708-65
5	39.6	3.2	14561	4	US-09-392-714-1
6	39.6	3.2	19124	2	US-08-487-826B-13
7	39.4	3.2	1664976	4	US-08-916-421B-1
8	39	3.2	10619	4	US-10-204-708-3
9	38.6	3.2	5511	3	US-08-928-361B-2
10	38.6	3.2	5511	4	US-09-588-995A-2
11	38.6	3.2	5526	3	US-08-751-359-21
12	38.6	3.2	5526	4	US-08-907-146-21
13	38.6	3.2	7334	3	US-08-928-361B-1
14	38.6	3.2	7334	4	US-09-588-995A-1
15	38.4	3.1	6801	4	US-10-204-708-61
16	38.4	3.1	13868	4	US-08-956-171B-173
17	37.8	3.1	640681	4	US-09-790-988-1
18	37.4	3.1	6659	4	US-10-204-708-6
19	37.2	3.0	11131	4	US-10-204-708-28
20	36.2	3.0	658	3	US-08-998-416-595
21	36.2	3.0	5562	4	US-10-204-708-63
22	36.2	3.0	10619	4	US-10-204-708-4
23	36.2	3.0	70000	4	US-09-851-896-3
24	36	2.9	240	1	US-08-628-417-6
25	36	2.9	399	4	US-09-621-976-8976
26	36	2.9	505	4	US-09-621-976-15639
27	35.8	2.9	6113	4	US-10-204-708-14

Sequence 20, Appl  
Sequence 57, Appl  
Sequence 39, Appl  
Sequence 26, Appl  
Sequence 12, Appl  
Sequence 43, Appl  
Sequence 16140, A  
Sequence 2813, A  
Sequence 424, Appl  
Sequence 1, Appl  
Sequence 84, Appl  
Sequence 93, Appl  
Sequence 54, Appl  
Sequence 85, Appl  
Sequence 1, Appl  
Sequence 3817, Ap  
Sequence 1, Appl

ALIGNMENTS

RESULT 1  
US-08-232-463-14  
; Sequence 14, Application US/08232463  
; Patent No. 5670367  
; GENERAL INFORMATION:  
; APPLICANT: DORNER, F.  
; APPLICANT: SCHEIFLINGER, F.  
; TITLE OF INVENTION: RECOMBINANT FOWLPOX VIRUS  
; NUMBER OF SEQUENCES: 52  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Foley & Lardner  
; STREET: 1800 Diagonal Road, Suite 500  
; CITY: Alexandria  
; STATE: VA  
; COUNTRY: USA  
; ZIP: 22313-0299  
; COMPUTER READABLE FORM: disk  
; MEDIUM TYPE: Floppy disk  
; OPERATING SYSTEM: IBM PC compatible  
; SOFTWARE: Patent In Release #1.0, Version #1.25  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/232,463  
; FILING DATE:  
; CLASSIFICATION: 435  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US/07/935,313  
; FILING DATE:  
; APPLICATION NUMBER: EP 91 114 300.6  
; FILING DATE: 26-AUG-1991  
; ATTORNEY/AGENT INFORMATION:  
; NAME: BENT, Stephen A.  
; REGISTRATION NUMBER: 29,768  
; REFERENCE/DOCKET NUMBER: 30472/114 IMMU  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (703)836-9300  
; TELEFAX: (703)683-4109  
; TELEX: 899149  
; INFORMATION FOR SEQ ID NO: 14:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 7218 base pairs  
; TYPE: nucleic acid  
; STRANDEDNESS: single  
; TOPOLOGY: linear  
; IMMEDIATE SOURCE:  
; CLONE: PTZ9pt-Fls  
US-08-232-463-14

Query Match 4.0%; Score 48.6; DB 1; Length 7218;

Best Local Similarity 5.1%; Pred. No. 0.0003;  
Matches 21; Conservative 217; Mismatches 171; Indels 0; Gaps 0;  
QY 56 TTAATCCTCATTAGGAACATGCGAGTCTTTGTTAACTCGCTCTACTTACCAATG 115  
Db 1028 TAAATCCGAGCTTGGCTGCGAGTGGAGGAGCTTGCATATTTTTTTTTTTTTTTTT 1087  
QY 116 GCGGAAACGAATCTGGTTTCATGAGCTTCCCTATCAACCGCTGGTTTCCCAATCATCTCA 175  
Db 1098 YY 1147  
QY 176 TCCCTCTCTTGGTCTCTCTCAGCGCTCGCGCGGCAACCGCAACCCCTAACACGCGG 235  
Db 1148 YY 1207  
QY 236 AAAACAAGCGGAACAAGCTTCTCCTCATGGAACCTCTCTGAGCGCCTCCATTTG 295  
Db 1208 YY 1267  
QY 296 TCATAGGGTGTCTACAGGACTTGACAATCTATATCTTACGATAGCATATCTGC 355  
Db 1268 YY 1327  
QY 356 CAGTTTCACTTCATCGCTCATATCGGAACCTCACTAGCTTCAACGCTCTCTCGCT 415  
Db 1328 YY 1387  
QY 416 TCTGTGTTAGTCAAGCAAGTTCACTCGCTTCTCATAAACGCGCTGT 464  
Db 1388 YYGT 1436

RESULT 2  
US-10-204-708-45  
; Sequence 45, Application US/10204708  
; Patent No. 667731  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication  
; TITLE OF INVENTION: by Assessing DNA Methylation  
; FILE REFERENCE: 5013.1012  
; CURRENT APPLICATION NUMBER: US/10/204,708  
; CURRENT FILING DATE: 2003-05-06  
; PRIOR APPLICATION NUMBER: PCT/EP01/03971  
; PRIOR FILING DATE: 2001-04-06  
; PRIOR APPLICATION NUMBER: DE 10019058.8  
; PRIOR FILING DATE: 2000-04-06  
; PRIOR APPLICATION NUMBER: DE 10019173.8  
; PRIOR FILING DATE: 2000-04-07  
; PRIOR APPLICATION NUMBER: DE 10032529.7  
; PRIOR FILING DATE: 2000-06-30  
; PRIOR APPLICATION NUMBER: DE 10043826.1  
; PRIOR FILING DATE: 2000-09-01  
; NUMBER OF SEQ ID NOS: 98  
; SEQ ID NO 45  
; LENGTH: 19233  
; TYPE: DNA  
; ORGANISM: Artificial Sequence  
; FEATURE:  
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)  
; NAME/KEY: unsure  
; LOCATION: (76, 178..179, 273, 586, 648, 651, 920, 1014, 1173, 1197, 1228)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1264, 1267, 1559, 1607, 1677, 1703, 1742, 1848, 1850, 2033)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (2168, 2254, 2257, 2280, 2291, 2723, 3018, 3080, 4205, 4209)

; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (4216, 4225, 4239, 4246, 4254, 4265, 4270, 4274, 4288, 4294..4295)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (4298, 4363, 4610, 4627, 4675, 4693, 4698, 4738, 4777, 4781)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (4787, 4825, 4829, 4838, 4855, 5462, 5494, 5496, 5500, 5674)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (5690, 5697, 5705, 5715, 5890, 5910, 5913, 6027, 6080, 6109)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (6244, 6288, 6776, 6873, 6891, 6980, 7028, 7656, 7876, 8030)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (8059, 8075, 8632, 8636, 9700, 9717, 9815, 9818, 9867, 9879)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1005, 10075, 10150, 10159, 10165, 10356, 11123, 11166, 11360)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (11703, 11933, 12050, 12340, 12981, 13564, 13764..13765, 13787)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (13846, 14168, 16566, 17006, 17217, 17650, 17653, 17656, 17898)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (18130, 18143, 18151, 18188, 18498, 18549, 18936, 18946, 19165)  
; OTHER INFORMATION: n is a or g or c or t  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (19200)  
; OTHER INFORMATION: n is a or g or c or t  
US-10-204-708-45  
Query Match 3.5%; Score 43; DB 4; Length 19233;  
Best Local Similarity 56.4%; Pred. No. 0.03;  
Matches 79; Conservative 0; Mismatches 61; Indels 0; Gaps 0;  
QY 1083 TTATGTTGCTTAATTTCTATAACTCTATACGATTATACAGACATTACTCTTTATGTTT 1142  
Db 10272 TTATTTGTTTAAAGTTGGTTTTTAATTTTGGTTTTTAAATAGTTTTTTTGTGTTT 10331  
QY 1143 GTTCTAAATATTATGCTGATTGCTGTTTTTCTGTTATTTCTGTTATAGTATGAA 1202  
Db 10332 TTAAGTTTTTGGATTATAGGTATGATGTTGTTGTTGTTGTTGTTTAAATAGAG 10391  
QY 1203 TAAATTTTGAAGATATTGA 1222  
Db 10392 TTATTTTGAATAGTTGA 10411

RESULT 3  
US-10-204-708-70  
; Sequence 70, Application US/10204708  
; Patent No. 667731  
; GENERAL INFORMATION:  
; APPLICANT: OLEK, Alexander  
; APPLICANT: PIEPENBROCK, Christian  
; APPLICANT: BERLIN, Kurt  
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication

```

; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 70
; LENGTH: 6040
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-70

Query Match      3.3%; Score 40; DB 4; Length 6040;
Best Local Similarity 57.08; Pred No. 0.12;
Matches 73; Conservative 0; Mismatches 55; Indels 0; Gaps 0;

QY 1083 TTATGTTGCTTAATTTCTATAAATTTCTATACGATTATAACAGAGCATCTACTGTTATGTTTT 1142
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4832 TTAGTTGGATGATTTATTCGATGGGAAGAATAATAGTAGATTTTAAATAACGTGTATG 4891

QY 1143 GTTCCTAAATATTAATGTCGATGTCGTTTTTTGTTATGTTCTGTCGTATAGTATGAA 1202
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4892 AGGAAGAAATTTATTTTAAATTCGCTGTCGTTGTTGTTTTTTATGTTGTTTATGTTGTA 4951

QY 1203 TAAAAATTT 1210
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 4952 TGGATTTT 4959

RESULT 4
US-10-204-708-65
; Sequence 65, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 65
; LENGTH: 5360
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-65

Query Match      3.2%; Score 39.8; DB 4; Length 5360;

```

Mon Sep 27 08:51:14 2004

us-09-913-767-1.rni

ADDRESSEE: Knobbe Martens Olson & Bear  
STREET: 620 Newport Center Drive 16th Floor  
CITY: Newport Beach  
STATE: California  
COUNTRY: US  
ZIP: 92660  
COMPUTER READABLE FORM:  
MEDIUM TYPE: Floppy disk  
COMPUTER: IBM PC compatible  
OPERATING SYSTEM: PC-DOS/MS-DOS  
SOFTWARE: PatentIn Release #1.0, Version #1.25  
CURRENT APPLICATION DATA:  
APPLICATION NUMBER: US/08/487,826B  
FILING DATE: 10-SEP-1993  
CLASSIFICATION: 435  
ATTORNEY/AGENT INFORMATION:  
NAME: Israelsen, Ned  
REGISTRATION NUMBER: 29,655  
REFERENCE/DOCKET NUMBER: NIH121.001CPI  
TELECOMMUNICATION INFORMATION:  
TELEPHONE: (619) 235-8550  
TELEFAX: (619) 235-0176  
INFORMATION FOR SEQ ID NO: 13:  
SEQUENCE CHARACTERISTICS:  
LENGTH: 19124 base pairs  
TYPE: nucleic acid  
STRADEDNESS: single  
TOPOLOGY: linear  
MOLECULE TYPE: CDNA  
HYPOTHETICAL: NO  
ANTI-SENSE: NO  
US-08-487-826B-13

Query Match 3.2%; Score 39.6; DB 2; Length 19124;  
Best Local Similarity 56.0%; Pred. No. 0.34;  
Matches 75; Conservative 0; Mismatches 59; Indels 0; Gaps 0;  
QY 1076 TTAGTGATTGTTGCTTAATTTCTATTAACCTTACGATATACAGACGATTTCTGTT 1135  
Db 15767 TTAATTTTTTTTATTTATGATATATATTTTATTTATTAATATATTTTTTCTTTTTT 15708  
QY 1136 ATGTTTGTCTTAATATATGATGATGTTGTTTGTATGTTCTTGTGATAA 1195  
Db 15707 TTGTTTTTATGATATATATTTTTTTTTTAAAGTTTTTTTTTCTTCTGTTT 15648  
QY 1196 GTATGAATAAAAT 1209  
Db 15647 ATTTTTTATAAT 15634

RESULT 7  
US-08-916-421B-1/c  
Sequence 1, Application US/08916421B  
Patent No. 6503729  
GENERAL INFORMATION:  
APPLICANT: Bult et al.  
TITLE OF INVENTION: Complete Genome Sequence of the Methanogenic Archaeon, Methanococcus jannaschii  
TITLE OF INVENTION: jannaschii  
FILE REFERENCE: PB275  
CURRENT APPLICATION NUMBER: US/08/916,421B  
CURRENT FILING DATE: 1997-08-22  
PRIOR FILING DATE: 1996-08-22  
NUMBER OF SEQ ID NOS: 3  
SOFTWARE: PatentIn version 3.1  
SEQ ID NO 1  
LENGTH: 1664976  
TYPE: DNA  
ORGANISM: Methanococcus jannaschii  
FEATURE:  
NAME/KEY: misc feature  
LOCATION: (28222)..(28222)

OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (28257)..(28258)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84773)..(84773)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84808)..(84808)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (84812)..(84812)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98120)..(98120)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98159)..(98159)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98239)..(98239)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98266)..(98266)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (98343)..(98343)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (103998)..(103998)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (148948)..(148948)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (163385)..(163385)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191989)..(191989)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (191995)..(191995)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (231980)..(231980)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234187)..(234187)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234220)..(234220)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (234814)..(234814)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309398)..(309398)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (309418)..(309418)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312837)..(312837)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (312993)..(312993)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (319226)..(319226)  
OTHER INFORMATION: n equals a, t, c, or g  
NAME/KEY: misc feature  
LOCATION: (559167)..(559167)  
OTHER INFORMATION: n equals a, t, c, or g



QY T078 TTAGTCAATTAGT

Query Match	3.2%	Score 39;	DB 4;	Length 10619;
Best Local Similarity	55.6%;	Pred. No. 0.36;		
Matches	75;	Conservative 0;	Mismatches 60;	Indels 0; Gaps 0;
Qy	1076	TTAGTGATTATGTTGCTTAATTTCTATAACTCTATACGATTATATAACAGAGCATTACTGTT	1135	



;; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF CHICKEN  
;; TITLE OF INVENTION: MONOCLONAL ANTIBODIES  
;; NUMBER OF SEQUENCES: 25  
;; CORRESPONDENCE ADDRESS:  
;; ADDRESSEE: Arnold, White & Durkee  
;; STREET: P.O. Box 4433  
;; CITY: Houston  
;; STATE: Texas  
;; COUNTRY: USA  
;; ZIP: 77210  
;; COMPUTER READABLE FORM:  
;; MEDIUM TYPE: Floppy disk  
;; COMPUTER: IBM PC compatible  
;; OPERATING SYSTEM: PC-DOS/MS-DOS  
;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/751,359  
;; FILING DATE: Concurrently Herewith  
;; CLASSIFICATION: 510  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Highlander, Steven L  
;; REGISTRATION NUMBER: 37,642  
;; REFERENCE/DOCKET NUMBER: ARSB:504  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 512/418-3000  
;; TELEFAX: 512/474-7577  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5526 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; US-08-751-359-21

Query Match 3.2%; Score 38.6; DB 3; Length 5526;  
Best Local Similarity 55.6%; Pred. No. 0.31; Mismatches 0; Gaps 0;  
Matches 74; Conservative 0

QY 1074 AGTTAGTGATTATGTTGCTTAATTTCTATAACTCTATACGATTATAACAGAGCACTACTG 1133  
DB 3334 ATTTGTCATTTTTTCGGAAATATATATATATATATATATATAAATAATTAATTTT 3393  
QY 1134 TTATGTTTGTCTCCTAAATATTATGTTGATGTTGTTTCTTGTATGTTGTTGTTAT 1193  
DB 3394 ATATATTTTATATATATATATATAAATAATATGTTATTTCTCTCTCTTTTATATAT 3453  
QY 1194 AGTATGAATAAA 1206  
DB 3454 ATATATAAATAATA 3466

## RESULT 12

US-08-907-146-21  
; Sequence 21, Application US/08907146  
; Patent No. 6316600  
; GENERAL INFORMATION:  
; APPLICANT: Michael, Nancy M  
; APPLICANT: Accavitti, Marianne  
; APPLICANT: Thompson, Craig B  
; TITLE OF INVENTION: METHODS FOR THE PRODUCTION OF CHICKEN  
; TITLE OF INVENTION: MONOCLONAL ANTIBODIES  
; NUMBER OF SEQUENCES: 25  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Arnold, White & Durkee  
; STREET: P.O. Box 4433  
; CITY: Houston  
; STATE: Texas  
; COUNTRY: USA  
; ZIP: 77210  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS

;; SOFTWARE: PatentIn Release #1.0, Version #1.30  
;; CURRENT APPLICATION DATA:  
;; APPLICATION NUMBER: US/08/907,146  
;; FILING DATE:  
;; CLASSIFICATION:  
;; PRIOR APPLICATION DATA: 08/751,359  
;; APPLICATION NUMBER:  
;; FILING DATE:  
;; ATTORNEY/AGENT INFORMATION:  
;; NAME: Highlander, Steven L  
;; REGISTRATION NUMBER: 37,642  
;; REFERENCE/DOCKET NUMBER: ARSB:504  
;; TELECOMMUNICATION INFORMATION:  
;; TELEPHONE: 512/418-3000  
;; TELEFAX: 512/474-7577  
;; INFORMATION FOR SEQ ID NO: 21:  
;; SEQUENCE CHARACTERISTICS:  
;; LENGTH: 5526 base pairs  
;; TYPE: nucleic acid  
;; STRANDEDNESS: single  
;; TOPOLOGY: linear  
;; US-08-907-146-21

Query Match 3.2%; Score 38.6; DB 4; Length 5526;  
Best Local Similarity 55.6%; Pred. No. 0.31; Mismatches 0; Gaps 0;  
Matches 74; Conservative 0

QY 1074 AGTTAGTGATTATGTTGCTTAATTTCTATAACTCTATACGATTATAACAGAGCACTACTG 1133  
DB 3334 ATTTGTCATTTTTTCGGAAATATATATATATATATATATATAAATAATTAATTTT 3393  
QY 1134 TTATGTTTGTCTCCTAAATATTATGTTGATGTTGTTTCTTGTATGTTGTTGTTAT 1193  
DB 3394 ATATATTTTATATATATATATAAATAATATGTTATTTCTCTCTCTTTTATATAT 3453  
QY 1194 AGTATGAATAAA 1206  
DB 3454 ATATATAAATAATA 3466

## RESULT 13

US-08-928-361B-1/c  
; Sequence 1, Application US/08928361B  
; Patent No. 6071518  
; GENERAL INFORMATION:  
; APPLICANT: Petersen, Carolyn  
; TITLE OF INVENTION: PEPTIDES, POLYPEPTIDES, GLYCOPROTEINS,  
; TITLE OF INVENTION: THEIR FUNCTIONAL MUTANTS, VARIANTS, ANALOGS AND FRAGMENTS  
; TITLE OF INVENTION: FOR TREATMENT AND DETECTION/DIAGNOSIS OF CRYPTOSPORIDIUM  
; TITLE OF INVENTION: SPECIES INFECTIONS  
; NUMBER OF SEQUENCES: 30  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: PETERS, VERNY, JONES & BIKSA  
; STREET: 385 Sherman Avenue, Suite 6  
; CITY: Palo Alto  
; STATE: CA  
; COUNTRY: USA  
; ZIP: 94306-1840  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: PatentIn Release #1.0, Version #1.30  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/08/928,361B  
; FILING DATE: 12-SEP-1997  
; CLASSIFICATION:  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: US 60/026,062  
; FILING DATE: 13-SEP-1996  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Verny, Hana  
; REGISTRATION NUMBER: 30,518

[illegible]

```

RESULT 15
US-10-204-708-61
; Sequence 61, Application US/10204708
; Patent No. 6677731
; GENERAL INFORMATION:
; APPLICANT: OLEK, Alexander
; APPLICANT: PIEPENBROCK, Christian
; APPLICANT: BERLIN, Kurt
; TITLE OF INVENTION: Diagnosis of Diseases Associated with DNA Replication
; TITLE OF INVENTION: by Assessing DNA Methylation
; FILE REFERENCE: 5013.1012
; CURRENT APPLICATION NUMBER: US/10/204,708
; CURRENT FILING DATE: 2003-05-06
; PRIOR APPLICATION NUMBER: PCT/EP01/03971
; PRIOR FILING DATE: 2001-04-06
; PRIOR APPLICATION NUMBER: DE 10019058.8
; PRIOR FILING DATE: 2000-04-06
; PRIOR APPLICATION NUMBER: DE 10019173.8
; PRIOR FILING DATE: 2000-04-07
; PRIOR APPLICATION NUMBER: DE 10032529.7
; PRIOR FILING DATE: 2000-06-30
; PRIOR APPLICATION NUMBER: DE 10043826.1
; PRIOR FILING DATE: 2000-09-01
; NUMBER OF SEQ ID NOS: 98
; SEQ ID NO 61
; LENGTH: 6801
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: chemically treated genomic DNA (Homo sapiens)
US-10-204-708-61

```

	Query Match	3.1%	Score 38.4	DB 4	Length 6801	
	Best Local Similarity	58.1%	Pred. No. 0.41			
	Matches 86	Conservative 0	Mismatches 61	Indels 1	Gaps 1	
Qy	1076	TTAGTGATTATGTTGCTTAATTTCTATACTCTATACGATTATAACAGACGACATTACTGT-	1134			
Db	3723	TAAGTTAATTTTGTGTTGGTTTTTAAATTTGTTGAGATTATAGGTGTGAATTAATTGTG	3782			
Qy	1135	TATGTTTTGTTCCCTAAATATATATGTGTGATGTGTGTTTTTGTATTGTCTCTGTGTATTA	1194			
Db	3783	TTTAGTTTAGTTTAAATTTTAATGAAGTTTAGTTTTTTTTTTTTTTGTTGTTGTGTTTT	3842			
Qy	1195	AGTATGAATAAAATTTGAAAGATATTCGA	1222			
Db	3843	TGATGTTATGGAATTTGAAATAGTTTTTA	3870			

Search completed: September 25, 2004, 10:10:44  
Job time : 109 secs

	Query Match	26.6%	Score 326;	DB 13;	Length 994;
	Best Local Similarity	62.2%;	Pred. No. 7.4e-84;		
	Matches 570; Conservative	0;	Mismatches 330;	Indels 16;	Gaps 3
QY	2	AAACAGCAAGCAGCAAGCAAGATGAAGAAATGGTTTGATAATCATAAACTGTATTATCC	61		
Db	95	AAGNAGCAAAAGCAGCAGAACCATGAAGCGCCTTCTTCACAGCAAACTGTCTTAC	154		
QY	62	TCAGTATAGAAACATGTGGAGGTCTTTGTTAACTCGTCTCTACTTCACCAATCGCGAA	121		
Db	155	TCACCATCGGCACCTCGGGTGGGCCCTCTCGTCATCGCTCTCTACTTCTCCACGCGCGCC	214		
QY	122	AAGCAATCTGGTTTCATGAGTTCCTATCAACCGCTGGTTTTCCAAATCATCTCTCATCCCTC	181		

```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12733C.1
; US-10-424-599-30274

Query Match      24.2%; Score 296.4; DB 13; Length 1233;
Best Local Similarity 62.4%; Pred. No. 3.6e-75;
Matches 481; Conservative 0; Mismatches 287; Indels 3; Gaps 1;

215 ACCGGTCTGGCTCCAGCTTCCTCGAAACCGCTGGCTTCCCTCATGCTCCTCCGCC 274
182 TCTTTGCTCTCTCTCAGCGGTGCGCGGGCAACCGCAACCCCTAACAAACCGGAAACA 241
275 TCGCGCTCTCTCTCTCAGCGGTGCGCGGGCAACCGCTTCCCTCCCGCGGAAACCGC 326
242 AGCGGAAACAAAGCTTCTCTCATGAAACTCTCTGTGGAGCGCTCCCATTTGTCATAG 301
327 -----CAAAACCGAATTAATCTCAATGAAGCTCTCTCTCCCTCGCGCTCCACCTTCATCG 382
302 GGTGTCTCAGAGGACTTGACAACCTATATATCTTACGGAATTAGCATATCTGCCAGTTT 361
383 GAAATCTCAGCGGCTCGAGACTACTCTACGGCTACGGCGTGGAGGCTTCCGGTCT 442
362 CAATCTCATCGTCTATATCGGAATCAACTAGCTTTCAAGCGCTCTCTTGGCTTTCTTGT 421
443 CCATCTCGCGCTCTCATATCGGAACGCACTCTCGGTTCACGGCTTCACGGCTTCCGCGTTCCTC 502
422 TAGTCAAGCAAAAGTTCACTCCGTTCTCCATAAAGCGCGTCTGTTGTGACGGTTGGTA 481
503 TCGTGAGGCAAGTTTACGGGTACTCCGTAAAGCGCGTCTGTTTGTCTCACTGTCCGCG 562
482 TCGGATCTCTCGGTACACAGTGATGAGACAAACCGGCTTAAGGAGCAAGAAAGT 541
563 CCGGGGTTTGGCGCTTCACACACGAGAGACCGTCCCGCTGGCGAGTCCGTTAAGGAAT 622
542 ATGTGGTTGGGTTCTTGATGACTGTGTTGCAAGCTCTCTCTATGCTTTTATATACCGC 601
623 ATGTATGGGCTTTGTGATGACAGTATCGTTCGGCATTTGATGGAATTCATTTTACGCT 682
602 TCGTTGAGCTAACTTACAGAAAGTCGTCAAGAAATCACTTCCCATCTGCTTGAGTA 661
683 TGTGGAGTTGTTGATACAAATAAATAAAGAGCTCTTACTACTCTCTGTGATGGAGA 742
662 TTCAGATGTCATGTCCTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 721
743 TTCAGTTCGTTATGTCCTTCTCGGCACCTCTCTTTGCTCTCTGGAATGATCAACA 802
722 GAGATTTAAGGTGATAGCAAGAGAGCAAGAGAGTTCGAAGATTTGAGAGTCAAGTTT 781
803 ATGACTTTAAGGTGATTCGAGGAGAGCCAAATAATTTGAGCAGGAG---AAGGAAGTT 859
782 ACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATG 841
860 ACTATGCTGTTTGGTGGGAGTGCATATATATGAGCAGGCTTTTCTTGGGGGCGATTG 919
842 GGATTTGTTTGTGATCATCATCTAGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 901
920 GGGTTATATTTGGGCTGCTTCT-CACTAGATTTTGTGATGGGTTGTTGCTACCGG 978
902 TGACTGAAGTTTTCG 917
979 TAACGGAAGTTTGGC 994

```

```

RESULT 2
US-10-424-599-30274
; Sequence 30274, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou Yihua
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 30274
; LENGTH: 1233
; TYPE: DNA

```

```

; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_12733C.1
; US-10-424-599-30274

Query Match      24.2%; Score 296.4; DB 13; Length 1233;
Best Local Similarity 62.4%; Pred. No. 3.6e-75;
Matches 481; Conservative 0; Mismatches 287; Indels 3; Gaps 1;

256 CTCCTCTCATGGAACCTCCTCTGTGGAGCGCTCCTCATATGTCATAGGGTGTGTACACAGA 315
136 CTAATCTCTATCAAGCCCTCTCTCTCTCGCTCCGCCCTCATCGGAATCTCTCACCGC 195
316 CTTGACAACTACTATATTTTACGATTTAGCATATCTGCCAGTTTCAACTTTCACTCGTCTC 375
196 CTCGAGCACTACTCTTACCGCTCGCGGTGGTTCGCTTCCGCTTCCACCTTCTCTCTA 255
376 ATAATCGGAACCTCAACTAGCTTTCAACGCTCTCTTCGCTTTCTTGTAGTCAAGCAAAAG 435
256 ATCCAAAGCCTCCACCTCGCTTTCACCGCGTCTTCGCGCTTCCCTCTCGTCCGCCACAGG 315
436 TTCACTCCGTTCTCCATATAACCGCGTCTGTTTGTGAGGTTGGTATCGGATCTCTGCG 495
316 TTCAACGCTTCTCTCTATGCTCTTCTATGCTTTTATATACCGCTCGCTGCTGCTGCTGCT 375
496 TTACACAGTGATGGAGACAAACCGGCTAAGAGAGACAAAGAGATATGTGTTGGGTTTC 555
376 CTCGGTTCAGCGGGACCGCCCGCGGTGAGTCGAGTCGCGAGTACCGTATGTTGTTT 435
556 TTGATGACTGTGTTGCGAGCTCTTCTATGCTTTTATATACCGCTCGTTCGCTGCTGCTGCT 615
436 GTTATGATATCTTCGCGCTCGCGCTGTATGGGTTGCTTTTGGCGTTGATGAGTTGGTG 495
616 TACAAGAAAGCTGCTCAAGAAATCACTTCCCACTTGTGCTGAGATTCAGATTCAGATTCATG 675
496 TACAAAAGAGAGAGGAGCGGTATCAGTACTCTCTGTCATGAGATTCAGCTTGTCTTG 555
676 TGCCTTCTCTCTACTTTTCTGTCATTTGTCATGTCATGTCATGTCATGTCATGTCATGTCATG 735
556 TGCTTCTTCTCTACTTATCTGACCGTTTGAATGATTAATCAATTAATGATCAAGTTTCAAGGTG 615
736 ATAGCAAGAGAGAGAGAGGATTCAGATTTGAGAGGATCACTGTTTACTATGATTCATGATA 795
616 ATTCGCGAGAGAGAGAGATTTTAAAGCTTGGGAAACAAAG---TACTACGTTGCTTG 672
796 GTGATCACAGAAATAATGCGCAAGGTTTCTTCTTAGGAGCCATAGGATTTGTTTGT 855
673 GTTGGAGTGCAATAATGTCGAGTTTCTTCTTGGGAGCAATAGGGGTTATCTTTTGT 732
856 GCATCATCATAGCTTCTGTTGTTCTGATAGTGTCTGCTTCCGCTGACTGAAAGTTTTC 915
733 GCCTCGTCTTGTGTTGTCGGGTATTAATTTGCTGCTTTTCTTCCAGTACGGAAGTTTG 792
916 GCCGTGCTTGTTCGCGGAGAGTTTTCAGGACAGAGAAAGGTTCTCTCTACCTTTCTTCT 975
793 GCTGTTATTTGATACAAAGAGAGCTTTCATGACAGAGAAAGGGTGTCTTGGTCTCTCT 852
976 CTTTGGGATTTGCTCTTACTTCTACGGCGAGTTTAAATCCGCAAGAA 1026
853 CTTTGGGGTTTGTGTCCTATTTCTATGGAGAGATAAAACAGACAGGAA 903

```

```

RESULT 3
US-10-425-114-10516
; Sequence 10516, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E.
; APPLICANT: Tabaska, Jack E.
; APPLICANT: Cao, Yongwei

```

; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 73128  
; SEQ ID NO 10516  
; LENGTH: 1146  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: 700941791\_FLI  
US-10-425-114-10516

Query Match 24.1%; Score 295.8; DB 13; Length 1146;  
Best Local Similarity 62.4%; Pred. No. 5.2e-75;

Matches 481; Conservative 0; Mismatches 287; Indels 3; Gaps 1;

```
QY 256 CTCCTCTCATGGAACCTCTCTGTGGAGCGCTCCATTGTCTATAGGTTGCTCACAGGA 315
DB 57 CTAACTCTATCAAGCCCTCTCTCTCGCTCGCCCTCATCGGACTCCTCACCGGC 116
QY 316 CTGACAACTACTATATCTTACGGATTAGCATATCTGCCAGTTTCAACTTTCATCGCTC 375
DB 117 CTCGACGACTACTCTACGCTCGGGTGGCTCGCTTCGGTCTCCACTTCTCTCTA 176
QY 376 ATAACTCGGAACCTCACTAGCTTTCAACGCTCTCTTCGTTTCTTGTAGTCAAGCAAAAG 435
DB 177 ATCAAAGCCTCCACCTCGCTTTCACCGCGCTCTCGCTTCTCTCGTCCGCCACAGG 236
QY 436 TTCACCTCCCTCTCCATAACCGCGCTGTTTGTACCGTGTGGTATCGGATCCTTGGC 495
DB 237 TTCACGCCCTACTCCGCTCACTCCGCTGCTTCTCACCGCTCGCCGCTGTGGCT 296
QY 496 TTACACAGTGATGGAGACAAACCGCTTAAGGAGAGCAAGAAAGATATGTGGTGGTTC 555
DB 297 CTGGCTTCAGCGGGAGCGCCCGCGGTGAGTGCAGTCCGACATCGTATGTTT 356
QY 556 TTGATGACTGTGGTTCAGCTCTCTCTATGCTTTTATATACCGCTCGTGTAGCTAACT 615
DB 357 GTTATGATACTTGGCGCTCGCGGCTGTATGGTTCGTTTGGCTTGATGGAGTTGGTG 416
QY 616 TACAAGAAAGCTCGTCAAGAAATCACTTCCCACTTGTGCTTGATGATGATGGTCAAG 675
DB 417 TACAAAAGAGCAGGACGCGTATCACGTAATCTCTGCTCATGAGATTCAGCTTGTCTTG 476
QY 676 TGCTTGTGCTACTTTTCTGTGTCATTGGCATGTTTCATCGTTCGAGATTTTAAGGTG 735
DB 477 TGTCTTCTTGTACTCTTCTGACCGCTTGGAAATGATTAATCAATGACITTCAGAGTG 536
QY 736 ATAGCAAGAGAGCAAGAGAGTTTCAAGATTGGAGGATCAGTGTGTTTACTATGATTTGATA 795
DB 537 ATTCCGCGAGNAGCAAGAGATTTTAAGCTTGGGNAACAAG---TACTACGTTGTGTTG 593
QY 796 GTGATCAGAGGAATAATATGCGAAGGTTTCTTCTTAGAGGCCATAGGAGATTTGTTTTGT 855
DB 594 GTGTGAGTGCAATAATATGCGAGTTTCTTCTTGGAGCAATAGGGGTTATCTTTTGT 653
QY 856 GCATCATCACTAGCTTCTGCTGTCTGATAAGTGTCTGCTTCGCTGACTGAGTTTTC 915
DB 654 GCCTGCTTGTGTGTCGGGTATTAATATGCTGCTTCTTCCAGTGAAGGATTTTG 713
QY 916 GCGCTGTTGTTGTTCCGGGAGAGTTTTCAGGACAGAGAAAGGTGCTCTCTACTTCTTCT 975
DB 714 GCTGTTATGTTATACAAAGAGACTTTCATGACAGAAAGGGGTTGTTGTGCTCTCT 773
QY 976 CTTTGGGATTTGTCTCTTACTTCTACGCGAGTTTAAATCCGCGCAAGAAA 1026
DB 774 CTTTGGGGGTTTGTCTTCTTCTATGAGAGATAAACAAGACAGGGAA 824
```

RESULT 4

US-10-437-963-57539

; Sequence 57539, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 57539  
; LENGTH: 1173  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_59340C.1  
US-10-437-963-57539

Query Match 17.4%; Score 212.6; DB 17; Length 1173;  
Best Local Similarity 52.7%; Pred. No. 9.5e-51;

Matches 542; Conservative 0; Mismatches 469; Indels 18; Gaps 3;

```
QY 10 AGCAGCAAGNAGATGAAGATGTTTGTATATCAATAAATGATTTATCTCCTACTATA 69
DB 88 AGCAGCGTAGACTGTTTCAGGAGCCGCTCCTCGTCAACTTCGTCCTCATGGTGGTC 147
QY 70 GGAACATGTGGAGTCTCTTTGTTAACTCTCTACCTTACCAATTCGCGGAAACGAATC 129
DB 148 GGTCTGGCTGCGGCGCTCTCTCTCCGCGCTACTTCTCGCGCGGCAACCGAAG 207
QY 130 TGGTTTCATGAGCTTCTTATCAACCGCTGGTTTTCGAATCATCTCATCTCTCTTTGGTC 189
DB 208 TGGCTCTCCAGCTTCTCCAGACCGCGCTGCTGCTCGCGCTCTGCTCGCGCTCTGCTTC 267
QY 190 TCGTTCTCAGCGCTGCGCGGCAACCGCAACCCCTTAAACAACGCGGAAACAAGCG--- 245
DB 268 TCGTACTCTCAGCGCGCGCGCGCGAGTTGAGGACGACGCGCTCGCGCTCGCGCG 327
QY 246 --GAAAAACAAAGCTCTTCTCATGGAACCTCTCTGAGGAGCGCTCCATTTGTCATAGG 303
DB 328 GGGCCACACCGCTGTTCTCTCATGACGCCACCGCTCTCTGTCGCTCGCGCTCGCGCG 387
QY 304 TTGCTCAGAGGACTTGACAACCTATTATATTCTTACGGATTAGCATATCTGCCAGTTTCA 363
DB 388 CTCATGACCGCGCTCGACGACCTCTCTACGCTACGGCTTGGCTTACCTCCCGGTGTC 447
QY 364 ACTTCATCGCTCATTAATCGGAACCTCAACTAGCTTTCAACGCTCTCTTTCGTTTCTGTTA 423
DB 448 ACCTCTCCATCTCTCACTCTCCAGAGCTGGCGCTTCAACGCGCGCTTTCGCGCTGCTGTC 507
QY 424 GTCAAGCAAAAGTTCACTCGTTCTCCATAAAGCGCTGTTTGTGTCGTTGATC 483
DB 508 GTGCGCACGGGTTCACGGCGTTCGCGTGAACCGCTGCTGCTGCTGCTGCTGCGCGCG 567
QY 484 GGGATCCTTTGCGTTTACACAGTGTAGAGACAAACCGGCTAAGGAGAGCAAGAAAGATAT 543
DB 568 GCGATGCTGGGATGAACCGCGGGGACCGCGCGGGGGGTGTCGCGCGCGAGTAC 627
QY 544 GTGGTTGGGTCTTGTAGTACTGTGGTTGAGCTTCTCTATGCTTTTATATPACGCTC 603
DB 628 TCGCGCGGTTCGCCATGACGCTCGCGCGCGCGCGCTGTACGGGCTCGTGTCTCCCGTC 687
QY 604 GTTGAGCTAACTTACAAGAA-----AGCTCTCAAGNAATCACTTTCCCACTTGTG 654
DB 688 ATGGAGCTCAGCCAGCGCGCAACACGCGCGCGCGCGCTGCTACGACGCTCGTTC 747
```



Qy 655 CTTGAGATTGAGTGGTCAATGCTGCTGCTACTTTTCTGTGTCATTTGGCATTTGGCATTTG 714  
Db 748 ATGAGATGCGAGTCTGTCATCGGCTTCTGTCGACCGGCTTTCAGCGCGTGGCATGCTC 807  
Qy 715 ATCGTTGGAGATTTTAAGTGTAGTCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 774  
Db 808 GTCAACACAGATTTTCCAGCGCAATCCAGGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 867  
Qy 775 GTGTTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 834  
Db 868 GGC---TACTACTGCTCTGCGCGGATCGCGCGGATGATGATGATGATGATGATGATGAT 924  
Qy 835 GCCATAGGAGATTTGTTTGGATCATCATCATCATCATCATCATCATCATCATCATCATCAT 894  
Db 925 ACGATCGCGCGCATCTTCTACGGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 984  
Qy 895 CTTCCGCTGATGAGTATTTCCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 954  
Db 985 ATCCCGGTACCGAGGTGCTCGCGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 1044  
Qy 955 GGTGCTCTCTACTCTTCTTCTTGGGATTTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1014  
Db 1045 GCGTGGCGCTCGCTCTCTGCTATGCGGCTTGGGCTTGGGCTTGGGCTTGGGCTTGGG 1104  
Qy 1015 TCCGGCAAG 1023  
Db 1105 GCGGCCAAG 1113

## RESULT 5

US-10-437-963-82061/c  
; Sequence 82061, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 82061  
; LENGTH: 1513  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_81525C.1  
US-10-437-963-82061

Query Match 11.4%; Score 139.2; DB 17; Length 1513;  
Best Local Similarity 50.5%; Pred. No. 2.8e-29;  
Matches 391; Conservative 0; Mismatches 378; Indels 5; Gaps 2;

Qy 293 TTGTCATAGGTTGCTCAGGACTTGACAACTACTATATTTTACGGATTAGCATATC 352  
Db 1011 TTGTCCTGGGCTCATCTGCTGCGAGACATGATGATTTCTTATGCTTACTATATC 952  
Qy 353 TGCAGTTTCAACTTCATCGCTCATATCGGAACCACTAGCTTTTCAACGCTCTCTCG 412  
Db 951 TTCCGGTCTCAACATATTCGCTCATCTGCTGCTAGTCAGCTTGCCTTCAATGCTGCTCT 892  
Qy 413 CTTTCTTGTAGTCAAGAAAGTTCACTCCGTTCTCCATAAACCGCGCTGTTTGTGA 472  
Db 891 CATATTCTCAATGCTCAAAAATTCACCCCTCTGATTTTCAATTCGGTAGTCTCTCTTA 832  
Qy 473 CGGTTGGTATCGGGATCTTGGCTTACACAGATGATGAGACAAACCGGCTAAGGAGCA 532

Db 831 CGTTTTCTGCTTCACTTCTTGGAGTTGATGAAGATTTCTCAGGGAATCTACTAGTATATCAC 772  
Qy 533 AGAAGAGATATGTTGGTTCCTTGAATGATGATGATGATGATGATGATGATGATGATGATGAT 592  
Db 771 ATGGGAATGACATTTTGGGTTTCTCTTGACACTAGGGGATCAGCTACATATCTCGCTCA 712  
Qy 593 TATTACCGCTCGTTGAGCTAACTTACAAGAAAGCTCGTCAAGAAATCCTTTTCCCACTTG 652  
Db 711 TTCTCTCCCTGATGCAAGTCACATTTGAGAAGGTTATTAAAGGGAGACCTTCTCAGTTG 652  
Qy 653 TGCTTCAGATTTCAGATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712  
Db 651 TGTGTAACATGCAAGATATATACAGCTCTCGTGGCAACATTTGGCTTCTCTTGTGGGTTAT 592  
Qy 713 TCATCGTTGGAGATTTTAAGGTGATGCAAGAGAGAGAGAGAGAGAGAGAGAGAGAGAGAG 772  
Db 591 TTGCAAGTGGTGAATGGATGACTTTACAAGAGAGAGATGATGATGATGATGATGATGATGAT 535  
Qy 773 CAGGTTTTACTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832  
Db 534 AGCTGTCATATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 475  
Qy 833 GAGCCATAGGAGATTTGTTTGTGCTATCATCATCATCATCATCATCATCATCATCATCATCAT 892  
Db 474 GAGTGGTGGATTTGATCTTTTGTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 415  
Qy 893 TGCTTCGGGTGACTGAAGTTTTCGCGCTGTTTGTTCGGGAGAGAGTTTCAGGACAGAGA 952  
Db 414 CTCTACCCATCATTTCTGTTTCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTG 355  
Qy 953 AAGGTGCTCTCTACTTCTTCTTGGGGATTTGCTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1012  
Db 354 AGATTATAGTATGCTGATGCGCATTTGGGATTTATGTCATATGCGCCACCAATTAATG 295  
Qy 1013 AATCCGGCAAGAAAGTTGTTGATAAACTCAACCGCGGAGAGAGAGAGAGAGAGAGAGAG 1066  
Db 294 TTGATGCCAAGAAAG--GTAGAAAGACTACAGTCAGTGTAGAGAAACTTCTCTTA 243

## RESULT 6

US-10-424-599-102519  
; Sequence 102519, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 102519  
; LENGTH: 2214  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_63592C.1  
US-10-424-599-102519

Query Match 11.2%; Score 137.8; DB 13; Length 2214;  
Best Local Similarity 50.6%; Pred. No. 9.3e-29;  
Matches 359; Conservative 0; Mismatches 347; Indels 3; Gaps 1;

Qy 293 TTGTCATAGGTTGCTCAGGACTTGACAACTACTATATTTTACGGATTAGCATATC 352  
Db 638 TGGTCTTTGGAGTCTTAATTCGCTGACAAATATGATGATGATGATGATGATGATGATGATGAT 697  
Qy 353 TCCAGTTTCAACTTCATCGCTCATATCGGAACCTCAACTAGCTTTCAACGCTCTCTTCG 412



Db 698 TCTCGGCTTCTACTATTCCTGCTGATTTGTGTCATCAACAGTTAGCTTTTAAATGCAAGTTTCT 757  
Qy 413 CTTTCTTGTGTAGCAAGCAAAAGTTCACTCCGTTTCTCCATAAAGCGCGTGTGTTGTTGA 472  
Db 758 CATATTTTATCAATCTCTCAAAAGTTTCACTGCTTGTATTAATACTCTACAGTGGTTCTCA 817  
Qy 473 CGGTTGGTATCGGATCCTTGTGTTTACAGTGTATGAGCAAAACCGGCTTAAGGAGCA 532  
Db 818 CTTTATCTGCTGACCTCTTGTGTTTAAAGCAAGCAAGATGAACCATCTGTTTCTTCCA 877  
Qy 533 AGAAAGATGATGTGTTGGTTCTTGATGATCTGTTGTTGAGCTTCTTCTATGCTTTTA 592  
Db 878 AGGGAAGTACATATTTGTTTCTATGTACCTTGGAGCTTCTGAGTGTACTCTCTTT 937  
Qy 593 TATTACCGCTGTGTAGTAACTTCAAGAAAGCTCTCAAGAAATCACTTTCCCACTTG 652  
Db 938 TGCTTTCCCTCATGCACTGACCTTTGAGAAGGTTCTGAAGAGGAAACATTTTCTGTTG 997  
Qy 653 TCGTTGAGATTCAGATGGTCACTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 712  
Db 998 TTTTGGAAATGCAAAATCTACATCAATCGTTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1057  
Qy 713 TCATCGTTGGAGATTTTAAAGTGATAGCAAGCAAGAGAGTTCAAGATTTGGAGGAT 772  
Db 1058 TTGCAAGTGGGAATGGCGTACTTTGCTGAGAGAAATGAGAGGTTTTCAGA---AAGGAT 1114  
Qy 773 CAGTGTCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 832  
Db 1115 ATGTTGCTTATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 1174  
Qy 833 GAGCCATAGGATGTTGTTGTCATGATGATGATGATGATGATGATGATGATGATGATGATGAT 892  
Db 1175 GTGTTGTTGGCTTGTCTTCTAGTGTCTTCTCTACTCAATGTTTAAAGCACAGTTT 1234  
Qy 893 TGCTTCCGTTGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGATGAT 952  
Db 1235 CTTTACCGGTAATCTCTATGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCTGCT 1294  
Qy 953 AAGGTGCTCTCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 1001  
Db 1295 AGATAATTTCTATGCTTTTGGCTCTATGAGGTTTGGCTCTTATATTA 1343

## RESULT 7

US-10-424-599-97043  
; Sequence 97043, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 97043  
; LENGTH: 1417  
; TYPE: DNA  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_58641C.1  
US-10-424-599-97043

Query Match 10.5%; Score 128.4; DB 13; Length 1417;  
Best Local Similarity 49.6%; Pred. No. 3.8e-26;  
Matches 365; Conservative 0; Mismatches 361; Indels 10; Gaps 1;  
Qy 278 TGTGAGCGCTCCATTTGTCATAGGTTGCTCAGAGCTTGACACTTATATCTT 337  
Db 457 TGTGATTATCTCAATTTCAATCGGAGTTATGTTAGGATTTCAACACCTTTCTCTCAT 516

Qy 338 ACGGATTAGCATATCTGCCAGTTTCAACTTCATCGCTCAATAATCGGAATCTCAACTAGCTT 397  
Db 517 GGGGAATCTTACTTCCCGTTTCAACCTCAGCACTTTTGTCTATCTTCCCACTACTCT 576  
Qy 398 TCAACGCTCTCTTGGCTTCTTGTAGTCAAGAAAGTTTCACTCCGTTTCTCCATAAAG 457  
Db 577 TCAACCTTCTTCTCTGTCTCATATAGTGAAGCAAAAAATAAATCTTTTCAAGCTGAAT 636  
Qy 458 CCGTCTGTTTGTGAGCGTTGTTATCGGATCTCTCGGTTTACACAGTGTATGAGAGCAAAAC 517  
Db 637 GCGTGTATCTTCTACCCCTAAGCTCCATCTTATAGCTTTGGAATCTAGCCACGAGAGAC 696  
Qy 518 CGCTAAGGAGAGCAAGAAAGATGATGTTGGTTTCTTGTAGTGTGTTGCGAGCTC 577  
Db 697 CAAAAGGCTTAATCAAAAGAACTACTTCAATGGAATCTTTTTCACCATAGGGGAGGTT 756  
Qy 578 TTTCTATGCTTTTATATTTACCGCTGTTGAGCTAACTTACAAGAAAGCTCTGTCAGAAA 637  
Db 757 TGAATGTTGCTTGTACTTGGCACTAATGGAAGAAATTTACAAGAAAGGTGAAT----- 809  
Qy 638 TCACCTTCCCACTTGTGCTTGAATTCAGATGCTCATGTGCTGCTGCTGCTGCTGCTGCTGCT 697  
Db 810 ----TGCTACCAATGTGTGGAATGCAATGGAATGGAATGGAATGGAATGGAATGGAATGGA 866  
Qy 698 GTGTCAATGGCATGTTTCTGTTGGAGATTTTAAAGGTGATAGCAAGAGAGCAAGAGAGT 757  
Db 867 CCATAGTTGGCATGACATGGATGGTGGATTTTCTGAGATGAAGTAGAAGCAATGG 926  
Qy 758 TCAAGATTTGGAGATCAGTGTGTTTACTATGATGATGATGATGATGATGATGATGATGATGAT 817  
Db 927 TCTTTGCAAAAGGAAGCAGAGATTTATTTGGGTGACAGTGTGGAATGTTGTTGCTGCTGCT 986  
Qy 818 AAGTTTCTTCTTAGCAGCATAGGATGCTGTTTGTGTCATCATCACTAGCTTCTGCTG 877  
Db 987 AGTATGTTTCTAGGAACTGCAAGAAATGTAATCTCTGACATCTTCACTGAGTGGGGGA 1046  
Qy 878 TTCTGATAAGTGTCTTCTGCTTCCGTCACCTGAAGTTTTCGCCGTCGTTTGTTCGGGAGA 937  
Db 1047 TTTGTCATGACATTTTGTGTCATGATGTTTAGGAGTGTGCTGCTTCTCAGAGATG 1106  
Qy 938 AGTTTCAAGCAGAAAGTGTCTCTACTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 997  
Db 1107 CATTTGGTGTGCAAAAGCTGTGCTCACTTTTCTGCAATTTTGGGGTCTGTTCTTATG 1166  
Qy 998 TCTACGGCGAGTTAA 1013  
Db 1167 TTTATGGCATCTACAA 1182

## RESULT 8

US-09-938-842A-2212  
; Sequence 2212, Application US/09938842A  
; Patent No. US20020160378A1  
; GENERAL INFORMATION:  
; APPLICANT: Harper, Jeff  
; APPLICANT: Kreps, Joel  
; APPLICANT: Wang, Xun  
; APPLICANT: Zhu, Xun  
; TITLE OF INVENTION: STRESS-REGULATED GENES OF PLANTS, TRANSGENIC PLANTS CONTAINING  
; TITLE OF INVENTION: SAME, AND METHODS OF USE  
; FILE REFERENCE: SCRI1300-3  
; CURRENT APPLICATION NUMBER: US/09/938,842A  
; CURRENT FILING DATE: 2001-08-24  
; PRIOR APPLICATION NUMBER: US 60/227,866  
; PRIOR FILING DATE: 2000-08-24  
; PRIOR APPLICATION NUMBER: US 60/264,647  
; PRIOR FILING DATE: 2001-01-16  
; PRIOR APPLICATION NUMBER: US 60/300,111  
; PRIOR FILING DATE: 2001-06-22  
; NUMBER OF SEQ ID NOS: 5379  
; SEQ ID NO 2212  
; LENGTH: 3387



Db 1093 TCATTGTCTATCAGCACTACCTCGACGAAAGAGTTGAAGACTAGCCACACAAGTCCT 1152  
QY 1051 GAGACAGAACTGCTATTCTTCCAGTTAGTGA 1082  
Db 1153 GTAGGAGATCCTCATCTACTACTCTGCTGAGGA 1184

RESULT 10  
US-10-425-114-13687  
; Sequence 13687, Application US/10425114  
; Publication No. US20040034888A1  
; GENERAL INFORMATION:  
; APPLICANT: Liu, Jingdong  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Screen, Steven E  
; APPLICANT: Tabaska, Jack E  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53313)B  
; CURRENT APPLICATION NUMBER: US/10/425,114  
; NUMBER OF SEQ ID NOS: 2003-04-28  
; SEQ ID NO 13687  
; LENGTH: 1296  
; TYPE: DNA  
; ORGANISM: Zea mays  
; FEATURE:  
; OTHER INFORMATION: Clone ID: LIB143-024-D8\_FLI  
US-10-425-114-13687

Query Match 10.3%; Score 126.6; DB 13; Length 1296;  
Best Local Similarity 51.0%; Pred. No. 1.2e-25;  
Matches 374; Conservative 0; Mismatches 329; Indels 30; Gaps 2;

QY 14 GCAAGAGAGATGAAGATGGTTTGATATCAATAAATCTGATATTCCTCACTATAGGAA 73  
Db 65 GCAGCAAGACGATCGCCGCTCTGGTGGCGACCACTCGTGATGCTCGCGCTGGCG 124  
QY 74 CATGTGGAGTCTTTGTTAACTCGTCTTACTTCCCAATGCTATCTCTCTCTCTCTCT 133  
Db 125 TCACCGGTGGGAGCTCTCAGCGCTCTTACTTTAGCAAGGGCGACACCGGAGTGGC 184  
QY 134 TCATGAGCTTCTATCAACCGGTGGTTTCCCAATCACTCTCACTCTCTCTCTCTCTCT 193  
Db 185 TCTCGGATGGTCCAGACCGGTGGTGGCGCTGCTCTCCGCGAGTGGCGCTCTCT 244  
QY 194 TCCTCAGCGTGGCGCGGACCGCAACCTTAAACACCGGAAACAAAGCGGAAACAA 253  
Db 245 AGTCCGCGCGGAGCGCGACCGGAGCGCCCGCTGCT- 284  
QY 254 AGCTCTTCTCATGGAATCTCTGTGGAGCGCTCCATGTCATAGGGTGTCTCAGAG 313  
Db 285 -GCTTACCAGACCCAGCGCGGATCTCTGTGGCGCGCGGCTCGGGCTCATCGCG 343  
QY 314 GACTTGACAACTTATATCTTACGAGTATAGCATATCTGCCAGTTTCAACTTTCATCGC 373  
Db 344 GGTGGACAACTTGTCTACGCTGGGCTCGAGTCTTCCCGCTCTCACTCTCGCA 403  
QY 374 TCATATCGGAATCACTAGCTTTCAACGCTCTCTTCTGCTTTCTTTAGTCAAGCAA 433  
Db 404 TCTCATCTCCACGAGCTGGCTTCAACGCTTCTGCTGCTGCTGCTGCTGCTGCTGCTG 463  
QY 434 AGTTCACTCGTCTCCATAAAGCGCTGTTTGTGAGCGTGTGATCGGGATCCTTG 493  
Db 464 GGTGACGATGGGACGGTGAACCGCTGGCTGCTGAGCGTGGCGCGCTGCTGCTGCTG 523  
QY 494 CGTTACACAGTATGGAGCAAAACCGCTTAAGAGAGCAAGAAAGATATGTTGGCT 553  
Db 524 GCTGATGCTCTCTGACCGCGCGCGGCTTACAGGAGTCTAGTACTGCTGGCT 583  
QY 554 TCTTGATGACTGTGGTGGAGCTCTTCTCTATGCTTTTATATTACCGCTCTGTAGCTAA 613

## RESULT 11

US-10-437-963-72452  
; Sequence 72452, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 72452  
; LENGTH: 789  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_72830C.1  
US-10-437-963-72452

Query Match 10.0%; Score 122.4; DB 17; Length 789;  
Best Local Similarity 52.9%; Pred. No. 1.4e-24;  
Matches 307; Conservative 0; Mismatches 246; Indels 27; Gaps 1;

QY 47 TAACTGTATTATCTCTACTATAGGAACATGTGGAGTCTTTTGTAACTGCTCTACT 106  
Db 230 TCAACTCGGGATGCTGTGGGCACACCGGGGGCGATCTCTTAGCGGCTCTACT 289  
QY 107 TCACCAATGGCGGAAACGAATCTGGTTCATGAGTCTCTATCAACCGTGGTTTCCAA 166  
Db 290 TCAGCAAGGGCGGCGCACAGGAAGTGGCTCTCGCGGTGGCTCGAGACCGGGTGGCCG 349  
QY 167 TCATCTCTATCTCTCTTGGTCTCTTCTCAGCGCTCGCGCGCAACCGCAACCTTA 226  
Db 350 TGTGTGTGTGCGGTGCTGTGGGTCTTACACAGCGCGCGCGCGCGCGCGCGCGCG 409  
QY 227 ACAACGCGGAAACAAAGCGGAAACAAAGCTCTTCTCATGGAACCTCTCTGTGGAGCG 286  
Db 410 C-----GCTGTCTCAGCGCGCGGCTGCTGCTG 442  
QY 287 CTTCATTTGTATAGGGTGTCTCAGGAGTTCAGCACTACTATATTTCTTACGATTAG 346  
Db 443 CGGGCGCGGTGTCTGCGGCTCTCACCGGCGCGGAGCACTTCGTCTACGCTACGGCTCG 502  
QY 347 CATATCTGCCAGTTTCAACTTCATCGTCAATATCGGAATCAACTAGCTTTTCAACGCTC 406  
Db 503 CTACTCTGCGGTGTCCACTCTCGGCATCTCATCTCCAGGAGTCTGCGTTTACCGTCT 562  
QY 407 TCTTCGCTTTCTTGTAGTCAAGCAAAAGTTTCACTCCGTTCTTCCATAAACGCGCTGTTT 466

Db 563 TCCTTCGGGTGCTCATCGTGGGCGAGCGGCTCACCGGGCGACGCTGAACCGCGTGGCG 622

Qy 467 TGTTCACGGTTGGTATCGGGATCCTTGGGTTTACACAGTGAATGGAGACAAACCGGCTAAGG 526

Db 623 TGCTCACTATCGGCGCGTCTGGTGGCTCCACGGGTCCAAGGACCGACCGCGCGGG 682

Qy 527 AGACAGAAAGAGATATGTGGTGGTCTTGTAGTACTGTGGTGGAGCTCTTCTCTATG 586

Db 683 TGACCAACCGGAAGTACTGGAATGGGCTTCTTCCTCACCCCTGGCGCGCGCGCTGTACG 742

Qy 587 CTTTATATTAACCGTCTGCTGAGCTAACTTACAAGAAAGC 626

Db 743 GGCTCATCTCCGCTCGTGGAGCTCGGTAACAAGCAGC 782

RESULT 12

US-10-424-599-31364

; Sequence 31364, Application US/10424599

; Publication No. US20040031072A1

; GENERAL INFORMATION:

; APPLICANT: La Rosa Thomas J

; APPLICANT: Kovalic David K

; APPLICANT: Zhou Yihua

; APPLICANT: Cao Yongwei

; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With

; FILE REFERENCE: 38-21(53223)B

; CURRENT APPLICATION NUMBER: US/10/424,599

; CURRENT FILING DATE: 2003-04-28

; NUMBER OF SEQ ID NOS: 285684

; SEQ ID NO 31364

; LENGTH: 1373

; TYPE: DNA

; ORGANISM: Glycine max

; FEATURE:

; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_128327C.1

US-10-424-599-31364

Query Match 9.6%; Score 117.4; DB 13; Length 1373;

Best Local Similarity 47.9%; Pred. No. 6e-23;

Matches 371; Conservative 0; Mismatches 401; Indels 3; Gaps 1;

Qy 240 CAAAGCGAAACAAAGCTCTTCTCATGAAACTCTCTGTGGAGCGCTCCATGTGTCAT 299

Db 449 CATCAGTAAACCTTTCCACCCTCTCTGAACTTAAACAGTCTGCTCTTATATGTTTT 508

Qy 300 AGGGTTGTCTCAGGACTTGACAACTACTATATTTCTTACGGAATGAGATATGTCAGT 359

Db 509 GGGTTTCTTAAGTGTGCTGATAAACCCTCATGATGCTTATGCTATGCTTACCTCCCTGC 568

Qy 360 TTCAACTTCATCGCTCATTAATCGGAACCTCACTAGCTTTCAAGCTCTCTTCGCTTTCT 419

Db 569 ATCCACCGCTCACTCGTGGCTTCATCACCCTTGTGTTTCTGCACTCTTTGGATATCT 628

Qy 420 GTTAGTCAAGCAAAAGTTTCACTCCGTTCTCCATAAACCGCGTCTGTTTGTGACGGTTGG 479

Db 629 TCTTGTGAAAAACAAGTAGATGCTTCGATAGTGAATTCGGTTTTCATCAATCAATTCG 688

Qy 480 TATCGGGATCTTGGGTTTACAGTGAATGAGACAAACCGGTAAAGGAGACAAAGAAAGA 539

Db 689 ATTGACCATCATTTGCACTGCACTCGAGTTTCAGAGATTCAGACAGATATGCAACATCAGTGCAGATGA 748

Qy 540 GTATGTGGTGGTCTTTCATGACTGTGGTTCGAGCTCTTCTCTATGCTTTTATATATACC 599

Db 749 GTACATCATGGGAATTTGTGGGAGTGTTTAGATCTGCTCTTTCACGGGCTTATATTCG 808

Qy 600 GCTCGTTGAGCTAACTTACAAGAAAGCTCGTCAAGAAATCACTTTCCCACTTTGTCTTGA 659

Db 809 TCTCTCGAGCTTGCTTTTGTGAAGTTGCTTGAAGAAGATCCTTCATCGTTGTTCTGGA 868

Qy 660 GATTCAAGTGGTCAATGCGCTTCTGCTACTTTTCTGTGTCATTTGGCATGTTTCTATCGT 719

Db 869 GCAGCAAGTCAATGGTTTCTCTGTTTTCATTTCTGTTTACCACTGTAGGATGATATGAG 928

Qy 720 TGGAGATTTTAAGTGATAGCAAGCAAGAGAGTTTCAAGATTTGGAGATCACTGTT 779

Db 929 TGTGTATTTTCAAGGATGACATGAGGCTACCCTTTCAAAGTGGTAGAAGTCTTAA 988

Qy 780 TTACTATGCATTTGATAGTATGATCACAGGAATAATATGGAAGGTTTCTTCTTAGAGACCAT 839

Db 989 TTATCTTTGTTATCATTTGGGTGCAATCACTTTTCAGCTGGGGTTCTGGGGGCACTGC 1048

Qy 840 AGGATTTGTTTGTGCAATCATCACTAGCTTCTGGTGTCTTGATAAGTGTCTGCTTCC 899

Db 1049 TGTAAATTTTCTGGGCTCTACTGTGCTAGCA---GGTGTCTTAAATGCAGTAAGAACC 1105

Qy 900 GGTGACTGAAGTTTTCGCGTCTGTTTTCGGGAGAAAGTTTCAGGCAGAGAAAGGTGT 959

Db 1106 CATAACAAGCATTCAGCTGTTATATCTGCTAAAGGACCCCTATGAGTGGTTTCAAGATCCT 1165

Qy 960 CTCCTACTCTTTCTTTCTTTGGGATTTGCTCTTACTTCTACGGCGAGTTTAAA 1014

Db 1166 CTCCTAGTATCACCTTTTGGGATTTGGGATTTGGCTCATATATTTATGGAAGTTCTAAA 1220

RESULT 13

US-10-416-898-9

; Sequence 9, Application US/10416898

; Publication No. US20040172670A1

; GENERAL INFORMATION:

; APPLICANT: Yale University

; APPLICANT: Walker, Elisabeth

; APPLICANT: Dellaporta, Stephen

; TITLE OF INVENTION: MAIZE YELLOW STRIPE1 AND RELATED GENES

; FILE REFERENCE: 44574-5106-US

; CURRENT APPLICATION NUMBER: US/10/416,898

; CURRENT FILING DATE: 2003-05-16

; PRIOR APPLICATION NUMBER: PCT/US01/43101

; PRIOR FILING DATE: 2001-11-16

; PRIOR APPLICATION NUMBER: US 60/249,222

; PRIOR FILING DATE: 2000-11-16

; NUMBER OF SEQ ID NOS: 20

; SOFTWARE: PatentIn version 3.1

; SEQ ID NO 9

; LENGTH: 83698

; TYPE: DNA

; ORGANISM: Zea mays

US-10-416-898-9

Query Match 9.6%; Score 117.2; DB 17; Length 83698;

Best Local Similarity 48.1%; Pred. No. 1.1e-21;

Matches 363; Conservative 0; Mismatches 388; Indels 3; Gaps 1;

Qy 275 CTCGTGGAGCGCCTCCATTTGCTCATAGGTTGCTCACAGGACTTGACAACACTTATATT 334

Db 67442 CTCGTGTTTGATTTATCTTCTCTGTTTGTCTATTGTTTGTAGATAATTTTATACT 67501

Qy 335 CTTACGAGTATAGATATTCGCCAGTTTCAACTTCATCGCTCATTAATCGGAATCAACTAG 394

Db 67502 CTGTTGCACTTTTGTATCTCTCTCTTCACTTATCGATTCTATGCTTTCACAGTTAG 67561

Qy 395 CTTTCAAGCTCTCTCGCTTTCTGTTAGTCAAGCAAAAGTTTCACCTCCGTTCTCCATAA 454

Db 67562 CTTTCAATGGTGTCTTCTATTTATCATCAATTCACAGAAATCATCTGTTGATTTCT 67621

Qy 455 ACGCGCTGTTTGTGACGGTTCGGATTCCTTGGGATCTTCTGCTTACAGTATGAGNACA 514

Db 67622 TCTCAGTGTGTTTCTCTATCTCTGCTGTGTTGGTTCTTCTTGCAGATGATTCAAATA 67681

Qy 515 AACCGCTTAAGCAGGACGAAGAGATGTTGGTCTTCTGATGCTGTTGTCAG 574

Db 67682 GCCCATCAGGAGATTTCTAAAGTGGAGTTACTTGAATGGTGTCTGTCAGTTTGTCT 67741

Qy 575 CTCCTCTCTATGCTTTTATATATACCGCTCGTGTGAGCTAACTTACAAGAAAGCTCGTCAAG 634

Db 67742 CTCCTATCTATCTCTTTCAGCTCTCTCTTATGCAAGTTTCTTTCGAGAAGGTTCTCAAGA 67801

QY 635 AATCACTTCCCACTTGTCTTGAAGATTCAGATGGTCAATGCTGTCTGTCTACTTTTT 694  
DB 67802 GTGAGACTCTCTATGGTCTCGAGATGCAATCTATACGTGCTTGTGCTTCTTGTG 67861  
QY 695 TCTGTCTCATTGCGATGTTTCATCGTTGGAGATTTAAGGTAGTAGCAAGAGCAAGAG 754  
DB 67862 TAGCGGTATTCGGAATTTTTCGCAAGCGGGAAATGGATGTTTGTGAGTGTGAGATGGAAG 67921  
QY 755 AGTTCAAGATTGGAGGATCAGTGTCTTACTATGATGATGATGATGATGATGATGATGAT 814  
DB 67922 AGTTTCAG---GAAGGTCAAGTCAATTTATGTTTGTGATTTGTGCGGGGAGCGGTTTCGT 67978  
QY 815 GCGAAGGTTTCTTCTTAGGAGCATAGGAGATGTTTGTGTCATCATCATCATCATCATCAT 874  
DB 67979 GTCAATTTGGGTTGTGTAGGAGCGGTGTGCGCTTATTTCTTGTGTTCTTGTGTTTCAA 68038  
QY 875 GTCTTCTGATAAGTGTCTCTCCGCTGAGTCAAGTTTTCGCGGTGTTTGTGTTTCCGGG 934  
DB 68039 ACCTTATTTAGTCTCTCACTCAATTTGTTACGCTCTCGCGGCAATTTGCGGTGTTCCATG 68098  
QY 935 AGAAGTTTTCAGGCGAGAGAAAGTGTCTCTCTACTTCTTCTTCTTCTTCTTCTTCTTCT 994  
DB 68099 ACAAGCTGACTGAGGTTAAGATGTTGCGGATGCCATCGCTTCAAGAGTTCACGTTT 68158  
QY 995 ACTTCTACGCGAGTTTAAATCCGCGCAAGAAAGT 1028  
DB 68159 ATATACCAGAACTATCTTGTGACTTGAAGT 68192

## RESULT 14

US-10-437-963-91425  
; Sequence 91425, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 91425  
; LENGTH: 1177  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_90000C.1  
US-10-437-963-91425

Query Match 9.2%; Score 112.8; DB 17; Length 1177;  
Best Local Similarity 47.9%; Pred. No. 1.2e-21;  
Matches 387; Conservative 0; Mismatches 417; Indels 4; Gaps 2;  
QY 258 CTTCTCTCATGGAACCTCTCTGTGGAGCGCTCCATTTGTCATAGGGTTGCTCACAGGACT 317  
DB 330 CCTCTCTCTCCCGCGGCTCTCTGCGCGCGCGCGCTCTGCGCGGCTGTACGCCGT 389  
QY 318 TGACAACTACTTATTTTACCGATTAGCATATCTGCCAGTTTCACTTTCATCGCTCAT 377  
DB 390 GTCGTGCTTGTGTACGCGCTGCGGTCGCGAGGCGCTGCCCTGTCCACGCTGCTCGCTGCT 449  
QY 378 AATCGGACTCACTAGCTTTCAAGCTCTCTTGGCTTCTTCTTCTTCTTCTTCTTCTTCTTCT 437  
DB 450 GCTGGCGACGACGTGGCTTCAACCGCGGTGTTGCGGTTCTCTTCTTCTTCTTCTTCTTCTTCT 509

QY 438 CACTCCGTTTCCATAAACCGCGTCTGTTTGTGACGGTTGGTATCGGGATCCTTTCGTT 497  
DB 510 CACGCGTCTTCGCGCAACCGCGTCTGTTGCTCTCACTGTGAGGAGCGTGTCCGAGGAAG 569  
QY 498 ACACAGTGTAGGAGACAAACCGGCTAAGGAGAGCAAGAGAGATATGTGTGGTGTCTT 557  
DB 570 GCGGTGTAAAGACGCGTCCGGACCCGATGACCGTGGAGGCTGCGCATCTAGTCCGACGTCCA 629  
QY 558 GATGACTGTGTTTGCAGCTCTTCTCTATGCTTTTATATATACCGTCTCTTCTGAGCTAACTTA 617  
DB 630 GCAAGTGTCTACTCTACAGTCGCGAGATTCGGATATGATGGAGTGTGCGCGCTTCTAGG 689  
QY 618 CAAGAAGCTCTCTCAAGAAATCACTTTCCACTTGTCTGTGAGATTCAGATGTGTCATGTG 677  
DB 690 TGGAGAATTTGCCGCTTACGAGCCCTACGCGACGCTGATGACAGATGCGAGCGGTGATGGG 749  
QY 678 CTTTGTGCT-CTACTTTTCTGTTCTATTTGGCATGTTTCATCGTTGGAGATTTTAAAGTGA 736  
DB 750 CGCGCGGACACCGCGGCTGTGCTCGGCATGCGCATCAAGGGCGCTTTCCAGGCGG 809  
QY 737 TAGCAAGAGAAAGCAAGAGATTCAGATTTGGAGGATCAGTGTTTTACTATGCAATGATAG 796  
DB 810 TGGCGCGGAGAGCGCGGCTTCGGGCTCGCG---CGGCCAATCTACTACTCTGCTCTCG 866  
QY 797 TGATCAGAGATTAATATGCAAGGTTTCTTCTTAGAGCCATAGGATGTGTGTTTGTG 856  
DB 867 CTTGGGACGCGTGTGCTGCGAGCTCTCAACCTGGGCATCATGGGCTCATCACCTGCG 926  
QY 857 CATCATCACTAGCTTCTGTTGTTCTGATAGTGTCTTCTGCTTCCGGTGAAGTGTTCG 916  
DB 927 CGTCTGCTGTCTCGCGGCTCATGATCGCGTGTCTCTGCGCTCTCTCGAGGCTCTCG 986  
QY 917 CCGTCTGTTTGTTCGCGGAGAGTTTTCAGGACAGAAAGTGTCTCTCTACTTCTTCTTCTC 976  
DB 987 CCGTCTATCTTCTTCCAGGAGAGTTTCAGCGGACGAGAGGATCGGCTCGTCTCTCGC 1046  
QY 977 TTTGGGATTTGTCTCTTCTTCTTACGCGGAGTTTAAATCCGCAAGAAAGTGTGTTGATA 1036  
DB 1047 TCTGGGATTTGCGCTCTCTCTCTACGCGAGAAAGGCGCAGAGAAAGAGGAGCGCAGA 1106  
QY 1037 AACCTCAACGCGCGGAGACAGAACTGCC 1064  
DB 1107 AGATCGCGAGCGCGAGAGGAGGTGCG 1134

## RESULT 15

US-10-437-963-43860/c  
; Sequence 43860, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 43860  
; LENGTH: 2381  
; TYPE: DNA  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(2381)  
; OTHER INFORMATION: unsure at all n locations  
; FEATURE:

Mon Sep 27 08:51:14 2004

us-09-913-767-1.rnpb

OTHER INFORMATION: Clone ID: PAT\_MRT4530\_46977C.1  
US-10-437-963-43860

Query Match		9.0%;	Score 110.2;	DB 17;	Length 2381;
Best Local Similarity		47.8%;	Pred. NO. 1.1e-20;		
Matches 353;		Conservative 0;	Mismatches 383;	Indels 3;	Gaps 1;
QY	289	TCCATTGTCATAGGCTTCTC	CACAGGACTTGACAACTACTTATAT	TCTTACGGATTAGCA	348
DB	2024	TACATCGTCTGGGCTCAT	ATCGCCCGCAGACATGATGATAC	TGTGGCTCAAG	1965
QY	349	TATCTGCCAGTTTCAACTTC	ATCGTCAATATCGGAATCAACTAG	CTTTCAACGCTCTC	408
DB	1964	TACCTCCCGCTTCGACCT	TACTCGCTCATCTGTGCCAGCAG	CTCGGTTCAATGTCTC	1905
QY	409	TTTCGGCTTTCTTGTAGT	CAAGCAAAAGTTCACTCCGTTCT	CCATAAACGCCGTCGTTTG	468
DB	1904	TTCTCATACGTCTCAACT	CCCGAAGGTCACTCTCTGATAT	TCAACTCCGTCGTGCTG	1845
QY	469	TTGACGGTTGGTATCGGG	ATCCTTGGTTTACACAGTGATGG	AGACAAACCGGCTAAGGAG	528
DB	1844	CTCACCATGTCGCTTCG	CTCATCGGAGTCAGCAAGGAG	TCTCAGGGGTCACCGGCTC	1785
QY	529	AGCAAGAAAGATATGTG	GTGGTTCTTTGATGACTGTGG	TGCGACTTCTCTATGCT	588
DB	1784	TCGGGAGGGAAGTATCT	GTCTCGGTTTCGTGTGACGCT	GGGGCGCTCGTGACCTACTCG	1725
QY	589	TTTATATTACCGCTCG	TTCAGCTTAATCAAGAAAGCT	CGTCAAGAAATCACTTTCCCA	648
DB	1724	CTGATCCTCGGCTGATG	AGCTCACTTCGAGACCATCAT	CAAGACACACACCTTCTCA	1665
QY	649	CTTGTGCTTGAGATTCAG	ATGTCATGTCCTTGTGCTACT	TTTTTCTGTGTCATTGGC	708
DB	1664	GCCGTCTCAACATGCAG	ATCTACACGGCGCTCGTGG	CAGCGCGCGCTCGGTGGT	1605
QY	709	ATGTTTCATCGTTGGAG	ATTTAAGGTGATACAGAGAG	CAAGAGTTCAAGATTGGA	768
DB	1604	CTGTTCCGAGCGCGAG	TGGAGTGCCTGAGGGGGGAG	ATGACGCTTCAGTTCGGGG	1545
QY	769	GGATCAGTGTTTTACT	ATGATGATGATCACAGGAAT	ATATGGCAAGGTTTCTTC	828
DB	1544	CAGT---TCTCCTAC	CTGATGACGCTGCTGGG	CGCGCTGTCGTGGCAGGT	1488
QY	829	TTAGGAGCCATAGGAT	TGTTTGTGCAATCATCACT	AGCTTCTGGTTCTGTATAGT	888
DB	1487	ATCGGGTGTCTGGCT	CTCATCTCGAGGTGTCGG	CGCTCTTCTCAACGTGAT	1428
QY	889	GTTCGTCTCCGGTGAC	TGAAGTTTTCGGCGTCGTT	TTGTTTCGGGAGAGTTT	948
DB	1427	GTGTCACTCGCGT	CACTCCCGTTCTTCGGG	TGTGTTCCACGAGGAT	1368
QY	949	GAGAAAGGTGTCTCT	CTTCTTCTTTTGGGAT	TTTGTCTTCTTCTACGG	1008
DB	1367	GTGAGATCGTGGCA	TGCTGATTGCAATTTGGG	ATTTATTTCTGTTCT	1308
QY	1009	TTTAAATCCGCGCA	AGAAAG	1027	
DB	1307	TATCTAGATGGCA	AGAAAG	1289	

Search completed: September 25, 2004, 11:39:02  
Job time : 654 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2004, 07:22:58 ; Search time 3626 Seconds  
(without alignments)  
10088.575 Million cell updates/sec

Title: US-09-913-767-1

Perfect score: 1225

Sequence: 1 aaacagcagcagcaagaa.....aatttgaaagatttgagct 1225

Scoring table: IDENTITY\_NUC

Gapop 10.0 , Gapext 1.0

Searched: 27513289 seqs, 14931090276 residues

Total number of hits satisfying chosen parameters: 55026578

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database :

EST:\*

1: em\_estba:\*

2: em\_esthum:\*

3: em\_estin:\*

4: em\_estnu:\*

5: em\_estov:\*

6: em\_estpl:\*

7: em\_estro:\*

8: em\_htc:\*

9: gb\_est1:\*

10: gb\_est2:\*

11: gb\_hic:\*

12: gb\_est3:\*

13: gb\_est4:\*

14: gb\_est5:\*

15: em\_estfun:\*

16: em\_estom:\*

17: em\_ges\_hum:\*

18: em\_ges\_inv:\*

19: em\_ges\_pln:\*

20: em\_ges\_vrt:\*

21: em\_ges\_fun:\*

22: em\_ges\_mam:\*

23: em\_ges\_mus:\*

24: em\_ges\_pro:\*

25: em\_ges\_rod:\*

26: em\_ges\_pbg:\*

27: em\_ges\_vrl:\*

28: gb\_ges1:\*

29: gb\_ges2:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

# SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	596.6	48.7	645	9 AU238506	AU238506
2	399.6	32.6	784	28 BH601144	BH601144 BOMJF33TF
3	384	31.3	827	28 BZ515292	BZ515292 BOMQL20TR
4	378.8	30.9	438	9 AU229698	AU229698 AU229698

C	5	349.6	28.5	688	28	BH589732	BH589732 BOCES04TR
	6	322.8	26.4	589	14	H76984	H76984 17415 lmbd
	7	306	25.0	680	12	BG439880	BG439880 GA_Ea000
	8	295.2	24.1	986	14	CK243751	CK243751 EST727388
	9	293.2	23.9	985	14	CK257466	CK257466 EST741103
	10	289.6	23.6	1013	14	CK251896	CK251896 EST735533
	11	288	23.5	998	14	CK247918	CK247918 EST731555
	12	277.4	22.6	907	14	CK257887	CK257887 EST741524
	13	275.6	22.5	988	14	CK248243	CK248243 EST731880
	14	273.4	22.3	979	14	CK252005	CK252005 EST735642
C	15	263	21.5	868	14	CK245030	CK245030 EST728667
	16	261.2	21.3	395	28	CK457536	CK457536 SALK_1106
	17	260.8	21.3	929	14	CK249872	CK249872 EST733509
	18	260.8	21.3	930	14	CK253429	CK253429 EST737066
	19	260.8	21.3	942	14	CK249994	CK249994 EST733631
	20	258.6	21.1	932	14	CK256828	CK256828 EST740465
	21	258	21.1	722	12	BM436574	BM436574 VVA008B03
	22	258	21.1	908	14	CK247701	CK247701 EST731338
	23	255.2	20.8	761	28	BZ501745	BZ501745 BONFA70TR
	24	254.8	20.8	703	14	CB006438	CB006438 VVC033G11
	25	251.2	20.5	971	14	CK272858	CK272858 EST718936
	26	249.8	20.4	701	13	BO874152	BO874152 QG14120.Y
	27	249.4	20.4	948	14	CK257752	CK257752 EST741389
	28	248.2	20.3	928	14	CK245031	CK245031 EST728668
	29	247.8	20.2	616	13	BO873355	BO873355 QG1IN19.Y
C	30	246.6	20.1	732	13	BO988613	BO988613 QGF15G06
	31	246.2	20.1	902	14	CK262518	CK262518 EST708596
	32	245	20.0	899	14	CK256093	CK256093 EST739730
C	33	244	19.9	527	28	AQ958362	AQ958362 LERAX12TR
	34	241.2	19.7	894	14	CK247243	CK247243 EST730890
	35	239.6	19.6	882	14	CK258020	CK258020 EST741657
	36	239.6	19.6	889	14	CK251126	CK251126 EST734763
	37	238.4	19.5	892	14	CK246384	CK246384 EST730021
	38	237.6	19.4	880	14	CK257437	CK257437 EST741074
	39	235.2	19.2	827	14	CK249190	CK249190 EST732827
	40	233.6	19.1	832	14	CK259060	CK259060 EST742697
	41	232.8	19.0	665	14	CD714562	CD714562 VVB202D06
	42	231.6	18.9	668	14	CB008408	CB008408 VVC051C02
C	43	220.6	18.0	665	13	BO799268	BO799268 EST 1437
	44	217.6	17.8	781	14	CK247244	CK247244 EST730881
	45	213.6	17.4	770	12	BM408148	BM408148 EST582475

## ALIGNMENTS

RESULT 1  
AU238506  
LOCUS  
DEFINITION AU238506 RAPL17 Arabidopsis thaliana cDNA clone RAFL17-32-O17 5', linear EST 01-APR-2002  
645 bp mRNA  
mRNA sequence.  
ACCESSION AU238506  
VERSION AU238506.1 GI:19877675  
KEYWORDS EST.  
SOURCE Arabidopsis thaliana (thale cress)  
ORGANISM Arabidopsis thaliana  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
Rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.  
1 (bases 1 to 645)  
Seki, M., Narusaka, M., Ishida, J., Kamiya, A., Satou, M., Nakajima, M.,  
Akiyama, K., Enju, A., Oono, Y., Sakurai, T., Carninci, P., Kawai, J.,  
Itoh, M., Ishii, Y., Arakawa, T., Shibata, K., Shinagawa, A.,  
Muramatsu, M., Hayashizaki, Y. and Shinozaki, K.  
Large scale analysis of Arabidopsis full-length cDNA

## REFERENCE

## AUTHORS

## TITLE

## JOURNAL

## COMMENT

## CONTACT

## UNPUBLISHED (2002)

## Plant Functional Genomics Research Group

## RIKEN Genomic Sciences Center

## 3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan

## Tel: 81-298-36-4359

## Fax: 81-298-36-9060

## Email: mseki@rtc.riken.go.jp







```

QY 491 TTGGTTACAGTGTGAGACAAACCGGCTAAGGAGACAAAGAGATATGTGTTG 550
Db 626 TGGCCTTGCAATACGAGGAGAGAGGCGCGCTGAGCAAG---GAGTATGTGTTG 682
QY 551 GGTTCCTTGAGTGTGTTGTCAGCTCTCTCTATGCTTTTATATACCGCTCGTTGAGC 610
Db 683 GGTTCATGTGATTTGATTTGAGCTGTCTCTACGCTTTTATGCGGCTCGTTGAGC 742
QY 611 TAACCTACAGAAAGTCGTCAGAAATCACTTTCCACATTG 652
Db 743 TTACTTTACAGAAAGTCGACAGAGATCACTTTACACATTG 784

RESULT 3
BZ515292/c
LOCUS BOMQL20TR BO_2_3_KB Brassica oleracea genomic clone BOMQL20,
DEFINITION genomic survey sequence.
ACCESSION BZ515292
VERSION BZ515292.1 GI:27043827
SOURCE GSS.
ORGANISM Brassica oleracea
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.
Town,C.D., Van Aken,S., Utterback,T., Koo,H. and Fraser,C.M.
Whole genome shotgun sequencing of Brassica oleracea
Unpublished (2001)
Other_GSSs: BOMQL20TF
Contact: Chris Town
TIGR
9712 Medical Center Drive, Rockville, MD 20850, USA.
Tel: 301-838-3523
Fax: 301-838-0208
Email: cdtown@tigr.org
DNA is from a doubled haploid provided by Tom Osborn.
Seq primer: TR
Class: sheared ends.

FEATURES             Location/Qualifiers
     source            1..827
                     /organism="Brassica oleracea"
                     /mol_type="genomic DNA"
                     /strain="TO100DH3"
                     /db_xref="taxon:3712"
                     /clone="BOMQL20"
                     /clone_lib="BO_2_3_KB"
                     /note="Vector: pBOS1; Site 1: BstXI; 2-3 kb sheared
                     genomic DNA inserted into pBOS1 using BstXI linkers"

ORIGIN
Query Match          31.3%; Score 384; DB 28; Length 827;
Best Local Similarity 69.8%; Pred. No. 3.2e-76;
Matches 582; Conservative 0; Mismatches 210; Indels 42; Gaps 3;

QY 13 AGCAAGAAGATGAAGATGTTTGATATCAATAAAGTATATCTCTACTATAGGA 72
Db 795 AGCAAGAGAGATGAAGACTGTCTGTATATCAATAAAGTATATCTCTGCTATGGA 736
QY 73 ACATGTGGAGGTCCTTTGTTAACTCGTCTCTACTTCCCAATGGCGGAAACGAATCTGG 132
Db 735 AACTGGGAGGCGCTCTAATACGCGTCTCTACTTCCAGAAACGTTGGCAACGATCTGG 676
QY 133 TTCATGAGCTTCTATCAACCGTGTGTTTTCGAATCATCTCATCCCTCTCTTGGTCTCC 192
Db 675 TTCTCAAGCTTCTCCAAACCGCAGGCTGTCCCATCATCTCTTCCCTCTCTCTCTCC 616
QY 193 TTCTCAGCGCTCGCGCGACACCGACCCCTTAACACGCGGAAACAAACGCGGAAACA 252
Db 615 TTCTCTCGGCGGT-----AAAGAAAGCAACACGACT 586
QY 253 AAGCTCTTCTCATGGAAACTCTCTGTGGAGCGCCTCCATTGTGCATAGGGTTGCTCACA 312

```

```

Db 585 CCATTTTCTCATAAACCTCTCTGTTCTTGGCTTCTATTGTGATGTGCTGCTCACT 526
QY 313 GGACTTGAACAACCTATATATTCTTACGGATAGCATATCTGCGCATTTCAACTTCATCG 372
Db 525 GGGTTGACAAATTAATCTCTACGCTTACGGGTAGCTTACCTCCAGTTTCCACTTCTCT 466
QY 373 CTCAATAATCGGAACCTCAACTAGCTTTCAAGCTCTCTTCCGCTTTTCTTGTAGTCAAGCAA 432
Db 465 TTGGTCAATCTCTCCCAACTGGGCTTCACTGCTTTCTTGGCCTTTTTCATGGTGAAGCAA 406
QY 433 AAGTTCACTCCGTTCTCCATAAAGCGGCTGTTTGTGTGACGGTGTGGTATCGGATCCCTT 492
Db 405 AAGTTCAAGCCCATTTATATAAAGCGGCTGTTTGTCTTACTCTGGGTGCGGAATCCCTT 346
QY 493 GCCTTACACAGTATGATGAGACAAACCGCTTAAGGAGACAAAGAGATATGTGTTGGG 552
Db 345 GCCCTTACACCGACGTCACAGCTTGCCAAAGAGACACACAAAGAGATATATCGTTGGG 286
QY 553 TTCTTGATGACTGTGTTGTCAGCTCTTCTCTATGCTTTTATATACCGCTCGTTGAGCTA 612
Db 285 TTCTCTATGACCCCTTCTGCGGCTATTTCTACGCGTGTGTTGCTTGGAGCTC 226
QY 613 ACTTACAGAAAGCTCGTCAAGAAATCACTTTCCCACTTGTGCTGCTGAGATTCAGATGCTC 672
Db 225 ACTTACAGAAAGCTCGTCAAGAAATCACTTACACGCTTGTGCTCGAGATGCAGTTGGTC 166
QY 673 ATGTGCTTGTGCTACTTTTCTGTCGTCATTTGCGATGTCATCTCGGAGATTTTAAAG 732
Db 165 TTGTGCTTGTGCTGCCACTTGTCTGCTGCTGGGATGCTAGCTGATGGCGATTTCAAG 106
QY 733 GTGATAGCAAGAGACAAAGAGATTTCAAGATTTGGAGGAT---CAGTGTTTTACTATGCA 789
Db 105 GAGTT-----ACCAAGAGATTTTAACTTGGAGGCTCTACTACGTATTAATGTTG 55
QY 790 TTGATAGTATACACAGAAATATATGCGAAGGTTTCTTCTTAGAGCCCATAGG 843
Db 54 GTGGTTGTGTTCACGGCCATCATCTGCGAAGGCTTTTCTTCTGGGATCTATTGGG 1

RESULT 4
AU229698/c
LOCUS AU229698 RAF17 Arabidopsis thaliana cDNA clone RAF17-32-O17 3',
DEFINITION mRNA sequence.
ACCESSION AU229698
VERSION AU229698.1 GI:19798174
SOURCE EST.
ORGANISM Arabidopsis thaliana (thale cress)
Arabidopsis thaliana
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;
rosids; eurosids II; Brassicales; Brassicaceae; Arabidopsis.
1 (bases 1 to 438)
Seki,M., Narusaka,M., Ishida,J., Kamiya,A., Satou,M., Nakajima,M.,
Akiyama,K., Enju,A., Oono,Y., Sakurai,T., Carninci,P., Kawai,J.,
Itoh,M., Ishii,Y., Arakawa,T., Shibata,K., Shinagawa,A.,
Murnatsum,M., Hayashizaki,Y. and Shinozaki,K.
Large scale analysis of Arabidopsis full-length cDNA
Unpublished (2002)
Contact: Motoaki Seki
Plant Functional Genomics Research Group
RIKEN Genomic Sciences Center
3-1-1 Koyadai, Tsukuba, Ibaraki 305-0074, Japan
Tel: 81-298-36-4359
Fax: 81-298-36-9060
Email: msekita@riken.go.jp
An Arabidopsis full-length cDNA library was constructed essentially
as reported previously (Seki et al., 1998). cDNA cleaved with BamHI
and XhoI was ligated to modified lambda FIC-1 vector (Carninci et
al., submitted for publication) digested with BamHI and SalI. This
clone is in a modified pBluescript vector. Please visit our web
site (http://www.gsc.riken.go.jp/e/plant/index\_e.html) for further

```

9712 Medical Center Drive, Rockville, MD 20850, USA.  
Tel: 301-838-3523  
Fax: 301-838-0208  
Email: cdtown@tigr.org  
DNA is from a doubled haploid provided by Tom Osborn.  
Seq primer: TR  
Class: sheared ends.

FEATURES  
source  
1. .688  
/organism="Brassica oleracea"  
/mol\_type="genomic DNA"  
/strain="TO1000DH3"  
/db\_xref="taxon:3712"  
/clone="BOGES04"  
/clone\_lib="BOGE"  
/notes="Vector: PHOS1; Site 1: BatXI; 2-3 kb sheared  
genomic DNA inserted into PHOS1 using BatXI linkers"

ORIGIN  
Query Match 28.5%; Score 349.6; DB 28; Length 688;  
Best Local Similarity 71.3%; Pred. No. 1.9e-68;  
Matches 482; Conservative 0; Mismatches 179; Indels 15; Gaps 1;  
QY 16 AAGAAGAAGATGAAGATGGTTTGATAATCATATAAACTGTATTATCTCATCTAGTAACA 75  
Db 664 ATGATGAAGATGAAGCGGTTCTTGTAAATCATAACTGATTATATTGGCGATAGAAAC 605  
QY 76 TGTGGAGGTCCTTGTAACTCGTCTACTTCCCAATGGCGGAAACGAATCTGGTTC 135  
Db 604 TGTGAGGGCCCTTAATGACGCGTCTACTTCTAGAAACGGTGGCAACGAACTCTGGTTC 545  
QY 136 ATGAGCTTCCTATCAACCGCTGGTGTTCCTCAATCATCTCATCCCTCTCTTGGTCTCTTC 195  
Db 544 TCAAGCTTCCTCCAACTCAGGTTCGCCCATCATCTCTCCCTCTCTCTCTCTCTCTTC 485  
QY 196 CTCAGCGCTCGCGCGCAACCGCAACCTTAAACAACGGGAAACAAAGCGGAAACAAAG 255  
Db 484 CTCACCGCTC-----ACCGTAAACAAGAACATACGAAGACTACA 440  
QY 256 CTCTCTCTCATGGAACCTCTCTGTGGGCGCTTCAATGTCTATAGGGTTTGTCTACAGGA 315  
Db 439 CTCTCTCTCATCAACCTCTCTGTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 380  
QY 316 CTTCGACAACTACTTATATCTTACGGATTAGCATATCTGCCAGTTTCAACTCATCGCTC 375  
Db 379 TTGCAAACTACTCTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTACGCTTAC 320  
QY 376 ATAATCGAACTCAACTAGCTTTCAACGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 435  
Db 319 ATCATCTCTCTCAGCTAGCGCTTACCGCTCTCTCTCTCTCTCTCTCTCTCTCTCTCTCT 260  
QY 436 TTCACTCGCTTCTCCATAAACCGCTGCTTTTGTGTGACGGTTGGTATCGGGATCCTTGGC 495  
Db 259 TTCAGCGCTTCTACTATTAACCGCTGCTTTTGTGCTCACTCTCGGCGCGGAAGCTTGCT 200  
QY 496 TTACACAGTGTAGGAGCAAAACCGGCTTAAGGAGAGCAAGAGATGTGTGGTTC 555  
Db 199 CTTTACGGGAGCGGTGACAGCTCCCCCAAGGAGAGCGCAAGAGATATATAGTTGGGTTT 140  
QY 556 TTGATGACTGTGGTTCAGAGCTTCTCTATGCTTTTATATTACCGCTCTGTTGAGCTAACT 615  
Db 139 GCTATGACGGTGCCTGACGCTGTTCTCTACCGGTTTGTGTGCGCTGTGTGAGCTTACT 80  
QY 616 TACAAGAAAGCTCGTCAAGAAATCACTTTCACCTGTGTGTGAGATTCAGATGGTCTATG 675  
Db 79 TACAAGAAAGCTAGTCAGAGGATCAGCTACACGCTTGTGTGTGAGATTCAGATGGTCTG 20  
QY 676 TGCCTTGTGTCTACTT 691  
Db 19 TCGTGGTGGCACTT 4

RESULT 6

details. Location/Qualifiers  
1. 438  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/db\_xref="taxon:3702"  
/clone="RAFL17-32-017"  
/lab\_host="DH10B"  
/clone\_lib="RAFL17"  
/note="Site 1: BamHI; Site 2: SalI; Subtraction Library.  
The sequence was obtained from samples subjected to  
dehydration-treated (1, 2, 5, 10 and 24 hr) and  
rehydration-treated (1, 2, 5, 10, and 24 hr after  
dehydration treatment)"  
ORIGIN  
Query Match 30.9%; Score 378.8; DB 9; Length 438;  
Best Local Similarity 98.4%; Pred. No. 4.6e-75;  
Matches 435; Conservative 0; Mismatches 2; Indels 5; Gaps 5;  
QY 785 ATGCATTGATAGTATACAGGAATAATATGCAAGGTTTCTTCTTAGGAGCCATAGGA 844  
Db 438 ATGCATTGATAGTATACAGGAATAATATGCG-AGGTTTCTTCTTAGGAGCCATAGGA 380  
QY 845 TTGTGTTTGTGCATCATCACTAGCTTCTGGTGTCTGATAGTGTCTTCTCGGTGA 904  
Db 379 TTGTGTTTGTGCATCATCACTAGCTTCTGGTGTCTGATAGTGTCTTCTCGGTGA 320  
QY 905 CTGAGTTTTCGCCGCTGTTTTCGCCGGAAGTTTCAGGCAGAGAAAGTGTCTCTC 964  
Db 319 CTGAGTTTTCGCCGCTGTTTTCGCCGGAAGTTTCAGGCAGAGAAAGTGTCTCTC 260  
QY 965 TACTTCTTCTCTTTGGGATTTGTCTCTTACTTCTACGGCGAGTTTAAATCCGCAAGA 1024  
Db 259 TACTTCTTCTCTTTGGGATTTGTCTCTTACTTCTACGGCGAGTTTAAATCCGCAAGA 200  
QY 1025 AGTTGTTGATAAACCCTCAACCGCGGAGACAGAA-CTGCCATTATCTCCAGTTAGTAT 1083  
Db 199 AAGTTGTTGATAAACCCTCAACCGCGGAGACAGAACTGCCATTATCTTCTCTTGTAGTAT 140  
QY 1084 TAGTTGCTTAATTTCTATAACTCTATAGCATATTAACAGAGCAATTACTGTTATGTTTG 1143  
Db 139 TAGTTGCTTAATTTCTATAACTCTATAGCATATTAACAGAGCAATTACTGTTATGTTTG 80  
QY 1144 TTCTTAATATATATGTGATGTGTGTTTGTGTTATGTTCTTGTGATATAGTATGAT 1203  
Db 79 TTCTT-ATAATTATGTGATGTGTG-TTTTGTTATTTGTTGTTGTTATGATGAT 22  
QY 1204 AAAATTGAAGATATGAGCT 1225  
Db 21 -AAATTGAAAGATATTGAGCT 1

RESULT 5  
BH589732/c  
LOCUS BH589732 BOGE Brassica oleracea genomic clone BOGES04, genomic  
DEFINITION survey sequence.  
ACCESSION BH589732  
VERSION BH589732.1 GI:17842184  
KEYWORDS GSS.  
SOURCE Brassica oleracea  
ORGANISM Brassica oleracea  
Eukaryote; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
rosids; eurosids II; Brassicales; Brassicaceae; Brassica.  
1 (bases 1 to 688)  
Town, C.D., Van Aken, S., Uterback, T., Koo, H. and Fraser, C.M.  
Whole genome shotgun sequencing of Brassica oleracea  
Unpublished (2001)  
Other\_GSSs: BOGES04TF  
Contact: Chris Town  
TIGR

H76984  
LOCUS 17415 Lambda-PRL2 Arabidopsis thaliana cdna clone 200N2T7, mRNA linear EST 05-JAN-1998  
DEFINITION sequence.  
ACCESSION H76984.1 GI:1054235  
VERSION H76984  
KEYWORDS Arabidopsis thaliana (thale cress)  
SOURCE Arabidopsis thaliana  
ORGANISM Arabidopsis thaliana  
REFERENCE 1 (bases 1 to 589)  
AUTHORS Newman, T., deBruijn, F.J., Green, P., Keegstra, K., Kende, H., McIntosh, L., Ohlrogge, J., Raikhel, N., Somerville, S., Thomas, M., Retzel, E., and Somerville, C.  
TITLE Genes galore: a summary of methods for accessing results from large-scale partial sequencing of anonymous Arabidopsis cdna clones  
JOURNAL Plant Physiol. 106, 1241-1255 (1994)  
MEDLINE 95148729  
PUBMED 7846151  
COMMENT Contact: Thomas Newman  
MSU-DOE Plant Research Laboratory  
Michigan State University  
MSU-DOE-PRL, Michigan State University, Plant Biology Bldg., E. Lansing, MI  
Tel: 517-353-0854  
Fax: 517-353-9168  
Email: 22313tcn@ibm.ci.msu.edu  
Seq primer: T7 dye primer.  
FEATURES  
source  
1. 589  
/organism="Arabidopsis thaliana"  
/mol\_type="mRNA"  
/strain="var columbia"  
/db\_xref="taxon:3702"  
/clone="200N2T7"  
/notes="Vector: lambda Zip-Lox; Site 1: Sal; Site 2: Not; Lambda PRL2 is a cDNA library derived from equal quantities of 4 pools of mRNA. The mRNA sources were 1) 7 day germinated etiolated seedlings; 2) tissue culture grown roots; 3) staged plants half with 24 hour light cycle, half on 16 hr light, 8 hour dark- rosettes; 4) same plants as 3 but aerial tissue (stems, flowers and siliques). The vector is BRL's lambda Zip-Lox. The cdna inserts were directionally cloned with Sal-Not arms using oligo dt primed cdna."  
ORIGIN  
Query Match 26.4%; Score 322.8; DB 14; Length 589;  
Best Local Similarity 95.1%; Pred. No. 2.2e-62;  
Matches 366; Conservative 0; Mismatches 13; Indels 6; Gaps 3;  
QY 55 ATTATCTCACTATAGAACATGTGAGGTCCTTTGTAACTCGTCTCTACTTCCACCAAT 114  
DB 1 ATTATCTCACTATAGAACATGTGAGGTCCTTTGTAACTCGTCTCTACTTCCACCAAT 60  
QY 115 GCGGGAACAGATCTGGTTCATGAGCTTCCTATCAACCGCTGGTTTCCAATCATCTTC 174  
DB 61 GCGGGAACAGATCTGGTTCATGAGCTTCCTATCAACCGCTGGTTTCCAATCATCTTC 120  
QY 175 ATCCCTCTCTTGTGCTCTCTCTCCAGCGTCGCGGCAACCGCAACCCCTAACACGCG 234  
DB 121 ATCCCTCTCTTGTGCTCTCTCTCCAGCGTCGCGGCAACCGCAACCCCTAACACGCG 180  
QY 235 GAAAAACAGCGGAAAAACAAAGCTCTTCTCATGGAACCTCCTCTGTGGAGCGCTCCATT 294  
DB 181 GAAAAACAGCGGAAAAACAAAGCTCTTCTCATGGAACCTCCTCTGTGTATCGCTCCATT 240  
QY 295 GTCATAGGGTGTCTCAGAGCTTGACAACTACTTATTTCTTACGGATTAGCATATCTG 354  
DB 241 GTCATAGGGTGTCTCAGAGCTTGACAACTACTTATTTCTTACCG -TTAGCATATCTG 299

QY 355 CCAGTTTCAACTTCATCGCTCATAAATC-GGAACCTCAACTAGCTTTT-----CAACGCTCTCT 409  
DB 300 CCAGTTTCAACTTCATCGCTCATAAATCGGGAACCTCAACTAGCTTTTCAACGGTCTCTTT 359  
QY 410 TCCTTTCTTTCTTTAGTCAAGCAAAA 434  
DB 360 CGGNTTCTTCTTTAGTCAAGCAAAA 384  
RESULT 7  
LOCUS BG439880  
DEFINITION GA\_Ea0005C03f Gossypium arboreum 7-10 dpa fiber library Gossypium  
arboresum cdna clone GA\_Ea0005C03f, mRNA sequence.  
ACCESSION BG439880  
VERSION BG439880.1 GI:13349537  
KEYWORDS EST.  
SOURCE Gossypium arboreum  
ORGANISM Gossypium arboreum  
REFERENCE 1 (bases 1 to 680)  
AUTHORS Wing, R.A., Frisch, D., Yu, Y., Main, D., Rambo, T., Simmons, J., Henry, D., Wood, T.C., Leslie, A. and Wilkins, T.A.  
TITLE An integrated analysis of the genetics, development, and evolution of the cotton fiber  
JOURNAL Unpublished (2000)  
COMMENT Contact: Wing RA  
Clemson University Genomics Institute  
Clemson University  
100 Jordan Hall, Clemson, SC 29634, USA  
Tel: 864 656 7288  
Fax: 864 656 4293  
Email: rwing@clemson.edu  
Seq primer: TAATACGACTCCTATAGG  
High quality sequence stop: 621.  
FEATURES  
source  
1. 680  
/organism="Gossypium arboreum"  
/mol\_type="mRNA"  
/strain="AKA"  
/cultivar="8400"  
/db\_xref="taxon:29729"  
/clone="GA\_Ea0005C03f"  
/tissue\_type="fibers isolated from bolls harvested 7-10 dpa"  
/lab\_host="E. coli"  
/clone\_lib="Gossypium arboreum 7-10 dpa fiber library"  
/notes="Vector: pBK-CMV; Site\_1: EcoRI; Site\_2: XhoI"  
ORIGIN  
Query Match 25.0%; Score 306; DB 12; Length 680;  
Best Local Similarity 66.6%; Pred. No. 1.4e-58;  
Matches 454; Conservative 0; Mismatches 225; Indels 3; Gaps 1;  
QY 323 ACTACTTATATCTTACGGATTAGCATATCTGCGCTTTCTTGTAGTCAAGCAAAAGTTCAC 382  
DB 1 ACTACTTACTCTATGCTTATCTGCTCTCCCGCTTTCGACTCTTCTTGTATCATCG 60  
QY 383 GAACCTCAACTAGCTTTTCAACGCTCTCTGCTTTCTTGTAGTCAAGCAAAAGTTCAC 442  
DB 61 CGTCGAGTGTGCTTTTCAACGCGGGGTTCCTTCTGTTGTGGAACAAAGTTCAC 120  
QY 443 CGTTCTCCATAAACCGCGTGGTTTGTGACCGTTCGGTATCGGGATCTCTTCGTTACACA 502  
DB 121 CCTACTCCATAAACCGCGTGGTTTGTGACCATAGGGCTGCGTTCCTGCGTTTCATT 180  
QY 503 GTGATCGAGCAAAACCGCTAAGGAGCAAGAAAGATATGTTGGTTCGTTTCATGA 562  
DB 181 CGAGCAGCAGCGCGCTGAGATGAATCCAAACAGGAATATATTTGGGGTTCGTAATGA 240

QY 563 CTGTGGTTCAGCTCTCTCTATGCTTTTATATTAATACCGCTCGTTGAGCTAACTTACAAGA 622  
DB 241 CCTAGCTGCAGCGCGCTTTGTACCGATTTGTATTTGCTTTGGTGGAACTAACTGTAAGA 300  
QY 623 AAGCTTCGTCAGAAATCACTTTCCCACTTGTGCTTGAGATTCAGATGGTCATGTGCGCTTG 682  
DB 301 AGGCAAAAGCAAGAGATCAGCTACGCCCTTTGTGATGGAGATTCAGCTGGTGATGTTGG 360  
QY 693 CTGCTACTTTTCTGTGTGCTATCGCATGTTTCATCGTTGGAGATTTTAAAGTGATAGCAA 742  
DB 361 TTGCTACTGGCTTTTGCACCGTTGGAGTGTCTGCTCAACAATGATTTCAAGTGATCGAA 420  
QY 743 GAGAAGCAAGAGATTCAGAGATTCAGAGATCAGTGTGTTTACTATGCAATGATGATCA 802  
DB 421 AGAAGCAAGGAAATTCAGGCTAGGAGAAACAAA---ATACTATGTGGTTGTGATTTT 477  
QY 803 CAGGAATAATAGGCAAGGTTTCTTCTTAGGAGCATAGGATTTGTGTTGTGTCATCAT 862  
DB 478 CTGCAATAATAAATCAATGTTTCTTCTGGGAGCTATTTGGAGTAGTATTTTGTGCGTCAT 537  
QY 863 CACTAGCTTCTGTGTTCTGATAAGTGTCTGCTTCCGGTGACTGAAGTTTTCGCCGTCG 922  
DB 538 CAATGCTATCGGGTGTAGTGATAGCGGTTCTATTAACAGTTACAGAGATTTTGGCAATAT 597  
QY 923 TTTGTTTCCGGGAGAGTTTTCAGGCGAGAAAGGTGCTCTCTACTTCTTCTTCTTCTTGGG 982  
DB 598 TTTCTCTCAATGAGAAGTTTCAAGCGGAGAAAGGTGCTCTCTGCTCTCTCTGCTGTTGGA 657  
QY 983 GATTGTCTCTTACTTCTACGG 1004  
DB 658 GCTTCGATAACTACTTTTATGG 679

RESULT 8  
CK243751 986 bp mRNA linear EST 12-DEC-2003  
LOCUS EST727388 potato callus cDNA library normalized and full-length  
DEFINITION Solanum tuberosum cDNA clone POCAS40 5' end, mRNA sequence.  
CK243751  
VERSION CK243751.1 GI:39788651  
KEYWORDS EST.  
SOURCE Solanum tuberosum (potato)  
ORGANISM Solanum tuberosum

REFERENCE Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
AUTHORS Generation of ESTs from potato callus tissue  
TITLE Unpublished (2003)  
JOURNAL The Institute for Genomic Research  
COMMENT 9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org

Seq primer: ATT TAG GTG ACA CTA TAG.  
Location/Qualifiers  
1. .986  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POCAS40"  
/tissue\_type="callus"  
/lab\_host="DH10B-RonA"  
/clone\_lib="potato callus cDNA library, normalized and  
full-length"  
/note="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Solanum tuberosum var.  
Kennebec callus tissue grown on solid media."

FEATURES  
source

ORIGIN

Query Match 24.1%; Score 295.2; DB 14; Length 986;

Best Local Similarity 58.3%; Pred. No. 4e-56;  
Matches 589; Conservative 0; Mismatches 393; Indels 28; Gaps 3;  
QY 18 GAAGAAGATGAAGATGTTTGTATTAATCAATAAATGATATATCTCTCACTATAGGAACATG 77  
DB 4 GAAATGAAGAAGCTTATCTTCTAAATTTTCAACATATATATGCTTCAATTCGCAATG 63  
QY 78 TGGAGTCTCTTTAACTCGTCTCTACTTCAACAAATGGCGGAAACAAATCTGTTTCAAT 137  
DB 64 CGGTGGTCTTTAAATCTCTGACTCTATTTTCAATGCGGCGCAACAAATTTGGATACC 123  
QY 138 GAGCTTCTCTATCAACCGCTGTTTCCAAATCATCTCATCTCTCTCTCTCTCTCTCTCT 197  
DB 124 CAGCGTATTAACAATCTGTTGGTTGTCAAATATTTCTCATCTCTCTCTCTCTCTCTCT 183  
QY 198 CAGCGTCTCCCGCGCAACCGCAACCCCTAACAAACCGGAAACAAAGCT 257  
DB 184 CCAACCGCGAAATTTCAAG-----ACCGAGGCTAAAT 219  
QY 258 CTTCTCTCATGGAATCTCTCTCTGAGCGCTCCATTTGTCTATAGGTTGCTCAGAGACT 317  
DB 220 CGTCTTTATCAACGCAAGATTCATCGCATCTCGCGCTCGGAATCATCTCGCTCT 279  
QY 318 TGCAACTACTTATATCTTACGATAGCATATCTGCGAGTTTCAACTTCACTCTCTCTCT 377  
DB 280 CGATGGTACTTAACTCATGGGACCCCGGAAATTTACCGTTTCACTTTCACACTAAT 339  
QY 378 AATCGGAATCTCACTACTAGCTTTCAACGCTCTCTCTGCTTTCTTTGTTAGTCAACAAAGTT 437  
DB 340 CAAACGCGACTCAACTTGGCTTCACTGCGCTTTTGTGCTTATAGTTTAAACAGAAAT 399  
QY 438 CACTCTGCTCTCCATAAACCGCTCTGTTGTTGAGCGTTGTTATCGGATCCTTGGCT 497  
DB 400 GACAGGTATTCGACGAATTCGCTGTTTGTCTTATCGCGGAGCTGCGACTTTAGCTCT 459  
QY 498 ACACAGTATGGAGACAAACCGCTAAGGAGACAAAGAGATATGTGTTGGGTTCTT 557  
DB 460 CCGGCGAAACGGTGACCGCGCGCGGAGTCCACGAGGATTTATATGTTGGGTTTGT 519  
QY 558 GATGACTGTGTTGACGCTCTCTCTATGCTTTTATATTAACGCTCTGTTGAGCTAATT 617  
DB 520 GATGACGTTATCGTGGCTGTTTATATGATTTGATGTTGCGCTTAAATTTGATTTTA 579  
QY 618 CAAGAAAGCTCGTCAAGAAATCACTTCCCACTTGTGCTTGAGATTCAGATTCATGTG 677  
DB 580 TATGAAGCGGAAAGCAAGCTGTTTACTACATCAGATTTGAGATTCAGATGTTTGGG 639  
QY 678 CTTTGTGCTACTTTTCTGTTGCTATTTGGCATGTTTCATCGTTGGAGATTTTAAAGTGAT 737  
DB 640 CATTTCTGCTACTGTTTGTGCTATTTGCACTATTTGGAATGATTTCAACAAAGGATTTCCAGGCGAT 699  
QY 738 AGCAAGAGAAAGCAAGAGATTTCAAGATTTGAGGATCAGTGTGTTTACTATGCTATGATGT 797  
DB 700 TCCAAGGGAAGCAAGTCAATATGCAATTTGGAGAACTTAAGTAT---TATATGTTACTAGT 756  
QY 798 GATCAGGAATATATGGCAAGGTTTCTCTTAGGAGCCATAGGAGATTTGTTTGTGTC 857  
DB 757 ATGGTGTGCCATTTATTTGGCAATTTCCGCTCTCTGGGTTAGTTGAGTATTTTTTATTC 816  
QY 858 ATCATCACTAGCTTCTGCTGTTTCTGATAAGTGTCTGCTCCGCTGCTGAGTTTTCG 917  
DB 817 TTCTCTTTTACTCTCTGGAATTTATAGGCGCTTTTACTCTCTTACCGAAGTTTGGC 876  
QY 918 CGTCTGTTTGTTCGGGAGAGATTTCAAGGAGAGAAAGGTTGCTCTCTCTCTCTCTCTCT 977  
DB 877 TGTAAATTTTGTTCACGAAATTTCAAGCCGAAAGAGAGTAGTATTTTTCTAGCTCT 936  
QY 978 -TTGGGATTTGCTCTCTTACTTCTACCGCGAGTTTAAATTCGCGCAAGAAA 1026  
DB 937 ATGGGGGATTTGTTTCATATTTTATGTTGTAATGAAACAAGGGGAAGAA 986

RESULT 9

```
CK257466
LOCUS CK257466 985 bp mRNA linear EST 12-DEC-2003
DEFINITION EST741103 potato callus cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POCD557 5' end, mRNA sequence.
ACCESSION CK257466
VERSION 1
KEYWORDS EST
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum

REFERENCE
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1. 985
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCD557"
/tissue_type="callus"
/lab_host="DH10B-TonaA"
/clone_lib="potato callus cDNA library, normalized and
full-length"
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN
Query Match 23.9%; Score 293.2; DB 14; Length 985;
Best Local Similarity 58.2%; Pred. No. 1.1e-55;
Matches 570; Conservative 0; Mismatches 383; Indels 27; Gaps 2;

QY 18 GAAGAAGTGAAGAAGTGGTTGATATATCAATAAAGTATATCTCCTCACTATAGAAACATG 77
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
33 GAAATGAAGAGAGCTATCTTCCTTAATTTCAACATAATATGCTTCAATGCGCAATTG 92
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB TGGAGTGCCTTCTTAACTCGTCTCTACTTCCAAATGCGGGAACGAATCTGGTTCAT 137
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB CGGTGGTCCCTTAACTCTGCTACTTATTCATTCATGCGGCGCAACGAATTTGGATACC 152
QY 138 GAGCTTCTATCAACCGCTGGTTTCCAAATCCTCCTCATCCCTCTCTTGGTCTCCTTCCT 197
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 153 CAGCGTATTACAACCTGTTGGTTGTCCTCAATATTTCTATCCCTCTAGCCATAGCCTATTT 212
QY 198 CAGCGTGGCGGCGCAACCGCAACCTCAACACGCGGGAACCAACGCGGGAACCAAGCT 257
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 213 CCAACGCGGGAATTAACAGG-----ACCGAGGCTAAAT 248
QY 258 CTTCTCTGGAACCTCCTCTGTGGAGCGCTCCATGTGTCATAGGTTGCTCAGACT 317
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 249 CGTCTTATACACGCGCAAGAGTTTCATCGCATTCGCGCGTCCGAATCATCGTCTGCT 308
QY 318 TGACAACTACTTATTTCTTACGGATTAGCATATCTGCCAGTTTCAACTTTCATCGTCTCAT 377
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 309 CGATGGTTACTTAACTATGCGGGAACCGGGAATTAACCGTTTCACTTCGACATAAT 368
QY 378 AATCGGAATCAACTAGCTTTCAACGCTCTCTTCGCTTTCTTGTAGTCAAGCAAAAGTT 437
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 369 CAACGCGACTCACTTGGTTCCTACTGCGCTTTTGTGCTGCTTATAGTTAAACAGAAAT 428
QY 438 CACTCGTCTTCATAAACCGCTCGTTTGTGAGCGTTGGTATCGGATTCCTTCGGTT 497
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 429 GACAGCGTATTCAGCAATTCGCTCGTTTGTCTTATCGCGGAGCTGCGACTTTAGCTCT 488
```

```
QY 498 ACACAGTGTAGAGACAAACCGGCTAAGGAGAGCAAGAAAGATGATGCTGGTTCCTT 557
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
489 CCGGCGGAACGGTGACCGGCGCGCGAGTCCACGAAGATATATGTTGGGTTGT 548
QY 558 GATGACTGTGTTGACAGCTCTTCTATGCTTTTATATACCGTCTGTTGAGCTAACTTA 617
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
549 GATGAGCTTTATCGTGGCGGTATATATGAGATTGATGTTGCCGCTTAATTCAGTTGATTA 608
QY 618 CAAGAAGCTCGTCAAGAAATCACTTTCCACCTTGTGCTTGAGATTTCAGATGTCATGTG 677
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
609 TATGAAGCGCAAGCAAGCTGTTACTACACTACAGTATTGGAGATTTCAGATGTTTGGG 668
QY 678 CTTGCTGCTACTTTTTTCTGTCTCATTTGGCATGTTTCATCGTTGGAGATTTTAAAGTGAT 737
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
669 CATTTCTGCTACTGTTTTTTTGGCACTATTTGGAATGGTTATCAACAAGATTTCCAGGCGAT 728
QY 738 AGCAAGAGCAAGAGAGTTCAGAGTTGAGAGTCAAGTGTGTTTATCATATGATGATGATG 797
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
729 TCCAAGGGAAGCAAGTCAATATGCAATTTGAGAACTCTAAG---TATTATATGCTACTAGT 785
QY 798 GATCAGAGGAATAATATGCAAGGTTTCTTCTTAGGAGCCATAGGAGTTGTTGTTTGTGC 857
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
786 ATGCTGTGCCATTTATTTGGCAATTCGCGCTCTCGGGGTAGTTGGAGTTATTTTATTTC 845
QY 858 ATCATCACTAGCTTCTGTTGTTCTGATAAGTGTTCCTCCGTTCCGTTGACTGAACTTTTCGC 917
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
846 TTCCTCTTTACTCTCTGGAATTATAGCGCGCTTTTCTTCTCTGTTTACCGAGTTTTCGC 905
QY 918 CGTCTGTTTCTCCGCGAGAGTTTCAGGAGAGAAAGTGTCTCTCTACTTCTTCTCTCT 977
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
906 TGTAAATTTTGTTCACGAAAAATTTCAAGCCGAAAAAGGAGTAGCTATTTTTCTAGCTCT 965
QY 978 TTGGGAGTTTGTCTCTTACT 997
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
966 ATGGGAGTTTGTTCATATT 985
```

```
RESULT 10
LOCUS CK251896
DEFINITION CK251896 1013 bp mRNA linear EST 12-DEC-2003
Solanum tuberosum cDNA library, normalized and full-length
Solanum tuberosum cDNA clone POCBR03 5' end, mRNA sequence.
ACCESSION CK251896
VERSION 1
KEYWORDS EST
SOURCE Solanum tuberosum (potato)
ORGANISM Solanum tuberosum

REFERENCE
AUTHORS Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.
TITLE Generation of ESTs from potato callus tissue
JOURNAL Unpublished (2003)
COMMENT The Institute for Genomic Research
9712 Medical Center Dr, Rockville, MD 20850, USA
Email: potato-array@tigr.org
Clones can be requested from TIGR via potato@tigr.org
Seq primer: ATT TAG GTG ACA CTA TAG.
Location/Qualifiers
1. 1013
/organism="Solanum tuberosum"
/mol_type="mRNA"
/cultivar="Kennebec"
/db_xref="taxon:4113"
/clone="POCBR03"
/tissue_type="callus"
/lab_host="DH10B-TonaA"
/clone_lib="potato callus cDNA library, normalized and
full-length"
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN
Query Match 23.9%; Score 293.2; DB 14; Length 985;
Best Local Similarity 58.2%; Pred. No. 1.1e-55;
Matches 570; Conservative 0; Mismatches 383; Indels 27; Gaps 2;

QY 18 GAAGAAGTGAAGAAGTGGTTGATATATCAATAAAGTATATCTCCTCACTATAGAAACATG 77
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
33 GAAATGAAGAGAGCTATCTTCCTTAATTTCAACATAATATGCTTCAATGCGCAATTG 92
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB TGGAGTGCCTTCTTAACTCGTCTCTACTTCCAAATGCGGGAACGAATCTGGTTCAT 137
QY ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB CGGTGGTCCCTTAACTCTGCTACTTATTCATTCATGCGGCGCAACGAATTTGGATACC 152
QY 138 GAGCTTCTATCAACCGCTGGTTTCCAAATCCTCCTCATCCCTCTCTTGGTCTCCTTCCT 197
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 153 CAGCGTATTACAACCTGTTGGTTGTCCTCAATATTTCTATCCCTCTAGCCATAGCCTATTT 212
QY 198 CAGCGTGGCGGCGCAACCGCAACCTCAACACGCGGGAACCAACGCGGGAACCAAGCT 257
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 213 CCAACGCGGGAATTAACAGG-----ACCGAGGCTAAAT 248
QY 258 CTTCTCTGGAACCTCCTCTGTGGAGCGCTCCATGTGTCATAGGTTGCTCAGACT 317
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 249 CGTCTTATACACGCGCAAGAGTTTCATCGCATTCGCGCGTCCGAATCATCGTCTGCT 308
QY 318 TGACAACTACTTATTTCTTACGGATTAGCATATCTGCCAGTTTCAACTTTCATCGTCTCAT 377
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 309 CGATGGTTACTTAACTATGCGGGAACCGGGAATTAACCGTTTCACTTCGACATAAT 368
QY 378 AATCGGAATCAACTAGCTTTCAACGCTCTCTTCGCTTTCTTGTAGTCAAGCAAAAGTT 437
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 369 CAACGCGACTCACTTGGTTCCTACTGCGCTTTTGTGCTGCTTATAGTTAAACAGAAAT 428
QY 438 CACTCGTCTTCATAAACCGCTCGTTTGTGAGCGTTGGTATCGGATTCCTTCGGTT 497
DB ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| ||||| |||||
DB 429 GACAGCGTATTCAGCAATTCGCTCGTTTGTCTTATCGCGGAGCTGCGACTTTAGCTCT 488
```



```

361 CAACGCGACCTCAACTCGCTTCACTGCGCTTTTGTGCTGTATAGTTAAACAGAAATT 420
438 CACTCCGCTTCCATAAACCGCGCTGTTTGTGACCGTGGTATCGGGATCCTTGGGTT 497
421 GACAGCGTATTCGACGAATTCGGTGGTGTGTTGCTATCGCGGAGCTGCGACTTAGCTCT 480
498 ACACAGTGTAGGAGACAAACCGGTAAAGGAGAGACAGAAAGAGTATGTGGTGGTTCCT 557
481 CCGGCGAACCGGTGACCGCGCGCGGAGTCCACGAAGGATTTATATGTTGGGGTTGT 540
558 GATGACTGTGTTGCGAGCTCTCTCTATGCTTTTATATACCGCTCGTTCAGCTTACTTA 617
541 GATGACGTTTATCGTGGGTGTTATGAGTATGTTGCCGTTAATTGAGTTGATTTA 600
618 CAAGAAGCTCGTCAAGAAATCACTTTCCACACTTGTCTTGAGATTGAGATGTCATGTG 677
601 TATGAAGCGCAAGCAAGCTGTACTTACACTACAGTATTCGAGATTCAGATGTTTGGG 660
678 CTTGCTGCTACTTTTCTGCTGTCATGGGATGTTTCATCGTTGGAGATTTAAGGTGAT 737
661 CATTTCTGCTACTGTTTCTGCTGCTATTTGCACTATTTGCAATGTTTATCAACAAGGATTTCCAGCGAT 720
738 AGCAAGAGAGCAGAGAGTTCAGATTTGAGATTCAGTGTGTTTACTATGCTATGATGT 797
721 TCCAAGGGAAGCAAGTCAATATGCAATTGGAGAACTTAAG---TATTATATGTTACTAGT 777
798 GATCAGAGGAATATATGGCAAGTGTCTTCTTAGGAGCATTAGGGATGCTGTTTGTGC 857
778 ATGGTGTGCCATTTATTTGGCAATTCGCGCTCTGGGTTAGTTGGAGTTATTTTATTC 837
858 ATCATCAGTACTGCTGTTGCTGCTGTAAGTGTGCTGCTGCTGCTGCTGCTGCTGCTGCT 917
838 TTCTCTTCTACTCTCTGGAATTTAGCGCTTTTCTTCTGTTTACCGAAGTTTGGC 897
918 CFTGCTGTTGTTCCGGGAGAGTTTCAGGAGAGAAAGTGTCTCTACTTCTTCTCT 977
898 TGTAAATTTGTTCCAGAAATTTTCAAGCCGAAAGAGGAGTAGCTATTTTCTAGCTCT 957
978 TTGGGGATTT 987
958 ATGGGGGATTT 967

```

CK257887 907 bp mRNA linear EST 12-DEC-2003  
 EST741524 potato callus cdna library, normalized and full-length  
 Solanum tuberosum cdna clone POCDB85 5' end, mRNA sequence.

CK257887  
 CK257887.1 GI:39814867  
 EST.  
 Solanum tuberosum (potato)  
 Solanum tuberosum

Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
 Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
 asterids; lamids; Solanales; Solanaceae; Solanum.

1 (bases 1 to 907)  
 Buell, C.R., Hart, A., Zismann, V., Karamycheva, S.A. and Baker, B.  
 Generation of ESTs from potato callus tissue  
 Unpublished (2003)  
 Contact: Robin Buell

The Institute for Genomic Research  
 9712 Medical Center Dr, Rockville, MD 20850, USA  
 Email: potato-array@tigr.org  
 Clones can be requested from TIGR via potato@tigr.org  
 Seq primer: Art TAG GTG ACA CTA TAG.

Location/Qualifiers  
 1..907  
 /organism="Solanum tuberosum"  
 /mol\_type="mRNA"  
 /cultivar="Kennebec"  
 /db\_xref="taxon:4113"

FEATURES  
 source

```

/clone="POCDB85"
/tissue_type="callus"
/lab_host="DH10B-Tona"
/clone_lib="potato callus cdna library, normalized and
full-length"
/notes=vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;
supplier: RNA was isolated from Solanum tuberosum var.
Kennebec callus tissue grown on solid media."

ORIGIN

Query Match      22.6%; Score 277.4; DB 14; Length 907;
Best Local Similarity 58.2%; Pred. No. 4.2e-52;
Matches 541; Conservative 0; Mismatches 361; Indels 27; Gaps 2;

QY 86 CTTTGTAACTCTCTCTACTTCCACCAATGCGCGAAACGAAATCTGTTTCATGAGCTTCC 145
DB 86 CTTTGTAACTCTCTCTACTTCCACCAATGCGCGAAACGAAATCTGTTTCATGAGCTTCC 145
QY 146 TATCAACCGCTGTTTTCATCATCTCTCATCCCTCTCTTGGTCTCTTCTCTCAGCCGTC 205
DB 146 TATCAACCGCTGTTTTCATCATCTCTCATCCCTCTCTTGGTCTCTTCTCTCAGCCGTC 205
QY 64 TACAAACTGTTGGTGTGTCCAATATTCTCATCCCTCTAGCCA TAGCTATTTCACAGCC 123
DB 64 TACAAACTGTTGGTGTGTCCAATATTCTCATCCCTCTAGCCA TAGCTATTTCACAGCC 123
QY 206 GCCGCGCAACCCCAACCTTAACACGCGGAAACGCGGAAACAAAGCTCTTCTCTCA 265
DB 206 GCCGCGCAACCCCAACCTTAACACGCGGAAACGCGGAAACAAAGCTCTTCTCTCA 265
QY 124 GGAATAATTCAGGACCCG-----AGGCTAAATCGTCTTTA 159
DB 124 GGAATAATTCAGGACCCG-----AGGCTAAATCGTCTTTA 159
QY 266 TGAATACTCTCTGTGAGCGCTCCATTTGTTCATAGGGTTGCTCACAGGACTTTCACAACT 325
DB 266 TGAATACTCTCTGTGAGCGCTCCATTTGTTCATAGGGTTGCTCACAGGACTTTCACAACT 325
QY 160 TCACACGGCAAGATTCATCGCATCTGCCGCGTCGAAATCATCGTGGTCTCGATGGTT 219
DB 160 TCACACGGCAAGATTCATCGCATCTGCCGCGTCGAAATCATCGTGGTCTCGATGGTT 219
QY 326 ACTTATATTCTTTACGATTAGCATATCTGCCAGTTTCAACTTCATCGCTCAATAATCGGAA 385
DB 326 ACTTATATTCTTTACGATTAGCATATCTGCCAGTTTCAACTTCATCGCTCAATAATCGGAA 385
QY 220 ACTTAAACTCATGGGACCGCGAAATTTACCGCTTTCACTTCGACACATAATCAACGCGA 279
DB 220 ACTTAAACTCATGGGACCGCGAAATTTACCGCTTTCACTTCGACACATAATCAACGCGA 279
QY 386 CTCAACTAGCTTTCAAGCTCTCTTCGCTTTCTTGTGTAGTCAAGCAAAAGTTCACTCCGT 445
DB 386 CTCAACTAGCTTTCAAGCTCTCTTCGCTTTCTTGTGTAGTCAAGCAAAAGTTCACTCCGT 445
QY 280 CTCAACTTGGCTTCACTGCGCTTTTGTGCTTATAGTTTAAACAGAAATTGACAGCGT 339
DB 280 CTCAACTTGGCTTCACTGCGCTTTTGTGCTTATAGTTTAAACAGAAATTGACAGCGT 339
QY 446 TCTCCATAAACGCGCTGTTTGTTCACGCTTGTATCGGATCCTTGGCTTACACAGTG 505
DB 446 TCTCCATAAACGCGCTGTTTGTTCACGCTTGTATCGGATCCTTGGCTTACACAGTG 505
QY 340 ATTGAGCAATTCGCTGCTTTGCTTATCGCGGAGCTGGAGCTTAGCTCTCGGGCGA 399
DB 340 ATTGAGCAATTCGCTGCTTTGCTTATCGCGGAGCTGGAGCTTAGCTCTCGGGCGA 399
QY 506 ATGGAGCAAAACCGGCTAAGGAGCAAGAAAGAGTATGTGGTGGGTTCTTGATGACTG 565
DB 506 ATGGAGCAAAACCGGCTAAGGAGCAAGAAAGAGTATGTGGTGGGTTCTTGATGACTG 565
QY 400 ACGGTGACCGCGCGCGGAGTCCACGAAGGATTATATGTTGGGGTTTGTGATGACGT 459
DB 400 ACGGTGACCGCGCGCGGAGTCCACGAAGGATTATATGTTGGGGTTTGTGATGACGT 459
QY 566 TGGTTCAGCTCTTCTCTATGCTTTTATATACCGCTCGTTGAGCTTAACCTTACAAGAAAG 625
DB 566 TGGTTCAGCTCTTCTCTATGCTTTTATATACCGCTCGTTGAGCTTAACCTTACAAGAAAG 625
QY 460 TTATCGGTGCGGTGTTTATATGATTTGATGTTGCGGTTAATTGAGTTGATTTATATGAAG 519
DB 460 TTATCGGTGCGGTGTTTATATGATTTGATGTTGCGGTTAATTGAGTTGATTTATATGAAG 519
QY 626 CTCGTCAAGAAATACCTTTCCCACTTGTGCTTGAGATTTCAGATGTCATGCTGCTGCTG 685
DB 626 CTCGTCAAGAAATACCTTTCCCACTTGTGCTTGAGATTTCAGATGTCATGCTGCTGCTG 685
QY 520 CGAAGCAAGCTGTACTTACCTACAGTATTTGAGATTTCAGATGTTTGGGCAATTTCTG 579
DB 520 CGAAGCAAGCTGTACTTACCTACAGTATTTGAGATTTCAGATGTTTGGGCAATTTCTG 579
QY 686 CTACTTTTTTCTGTGCTATTGTCATCTTCATCGTTGAGATTTTAAGGTTGATAGCAAGAG 745
DB 686 CTACTTTTTTCTGTGCTATTGTCATCTTCATCGTTGAGATTTTAAGGTTGATAGCAAGAG 745
QY 580 CTACTGTTTTTGGCACTATTGGAATGATTATCAACAAGGATTTCAGGCGGATTCGAAGG 639
DB 580 CTACTGTTTTTGGCACTATTGGAATGATTATCAACAAGGATTTCAGGCGGATTCGAAGG 639
QY 746 AGCAAGAGAGATTCAAGATTGAGGATCAGTGTGTTTATGCAATTTGATGATGATGATCAG 805
DB 746 AGCAAGAGAGATTCAAGATTGAGGATCAGTGTGTTTATGCAATTTGATGATGATGATCAG 805
QY 640 AGCAAGTCAATATGCAATTTGGAGAACTAAG---TATTATATGTTAGTATGTTGTTG 696
DB 640 AGCAAGTCAATATGCAATTTGGAGAACTAAG---TATTATATGTTAGTATGTTGTTG 696
QY 806 GAATATATGGCAAGGTTTCTTCTTAGGAGCCATAGGATTGTTGTTTGTGATCATCATC 865
DB 806 GAATATATGGCAAGGTTTCTTCTTAGGAGCCATAGGATTGTTGTTTGTGATCATCATC 865
QY 697 CCATTATTTGGCAATTTCCGCGCTCTCGGGGTAGTTGGAGTTATTTTATTTCTTCTCT 756
DB 697 CCATTATTTGGCAATTTCCGCGCTCTCGGGGTAGTTGGAGTTATTTTATTTCTTCTCT 756
QY 866 TAGCTTCTGCTGTTCTGTAAGTGTGTTCTGCTTCCGCTGAGCTGAAGTTTTCGCCGCTG 925
DB 866 TAGCTTCTGCTGTTCTGTAAGTGTGTTCTGCTTCCGCTGAGCTGAAGTTTTCGCCGCTG 925
QY 757 TACTCTCTGGAATTTATAGGCGCTTTTACTTCTCTGTTACCGAAGTTTGGCTGTAA 816
DB 757 TACTCTCTGGAATTTATAGGCGCTTTTACTTCTCTGTTACCGAAGTTTGGCTGTAA 816
QY 926 GTTTCGGGAGAGTTTTCAGGAGAGAAAGGTTGCTCTCTACTTCTTCTTCTTGGGAT 985
DB 926 GTTTCGGGAGAGTTTTCAGGAGAGAAAGGTTGCTCTCTACTTCTTCTTCTTGGGAT 985
QY 817 TGTTCACGAAATTTTCAAGCCGAAAGGAGTAGCTATTTTCTTAGCTCTATGCGGAT 876
DB 817 TGTTCACGAAATTTTCAAGCCGAAAGGAGTAGCTATTTTCTTAGCTCTATGCGGAT 876

```



QY	986	TTGTCCTTACTTCTACGCGAGTTTAA	1014	988 bp	mRNA	linear	EST 12-DEC-2003
Db	877	TTGTTTCATATTTTATGTTGAATGAA	905				Solanum tuberosum cDNA library, normalized and full-length
RESULT 13							
CK248243							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
COMMENT							
FEATURES							
source							
ORIGIN							
Query Match							
Best Local Similarity							
Matches							
QY	18	GAAGAAGATGAAGATGGTTTGAATATCAATAACTGTATTATCTTACCTCCTACATATAGGAACATG	77				
Db	39	GAATATGAAGAAGCTTATCTTCTTAATTTCAACATAATTATGTCTTCAATCGGCAATG	98				
QY	78	TGAGGTCTCTTGTGTTAACTCGTCTCTACCTTCAACATGCGGAAACGAATCTGGTTTCAT	137				
Db	99	CGTGTGCTTTTAACTCTCGACTCTATTTCATTATGCGGCGCAAGAAATTTGGATACC	158				
QY	138	GAGCTTCTTATCAACCGCTGGTTTTCATATFCTTCATCCCTCTCTTGGTCTCTCTTCT	197				
Db	159	CAGCGTATTACAACTGTTGGTTGTCCAAATTTCTCATCCCTCTAGCCATAGCCCTATT	218				
QY	198	CAGCCGTGCGCGGCAACCGCAACCTTAACACGCGGAAACAAACGAAGCT	257				
Db	219	CCAACCGCGGAAATTCAGG-----ACCGAGGCTTAAT	254				
QY	258	CTTCTCTCATGAAACTCTCTGTGGAGCGCTCCATTGTTCATAGGGTTGCTCACAGGACT	317				
Db	255	CGTCTTTATCACACGCGAAGATTTCATCGCATCTGCGGCGTGGGAATCATCGTGGTCT	314				
QY	318	TGACAACTACTTATATTTTACGGATTAGCATATCTGCCAGTTTCACTTCATCGTCTAT	377				
Db	315	CGATGGTTACTTAACTCATGGGACCCCGGAAATTAACCGTTTCACTTCGACTTAAT	374				
QY	378	AATCGGAAGCTCACTAGCTTTTCAACGCTCTCTTCGCTTCTTCTTAGTCAAGCAAGATT	437				
Db	375	CAACGGAGCTCAACTTTCGGTTCACTGCGCTTTTTCGCTTATAGTTAAACAGAAATT	434				
QY	438	CACCTCGCTTCTCCATAAACCGCTCGTTTGTGACGGTTGGTATCGGATCCCTCGGTT	497				
Db	435	GACAGCGTATTTCGACGAATTCGCTCGTTTATCGCGCGAGCTCGCACTTTAGCTCT	494				
QY	498	ACACAGTGTATGGAGACACAAACCGCTAAGGAGACGAAGAAGAGTATGTGTTGGTTCTT	557				
Db	495	CCGGCGGAACGGTGAACCGCGCGCGGAGTCCACGAAGATTATATGTGCGGTTTGT	554				
QY	558	GATGACTGTGGTTGACAGCTCTTCTCTATGCTTTTATATTTACCGCTCGTTGAGTAACTTA	617				
Db	555	GATGACGTTTATCGTGGTGTATATGATGATTGATGTTGCCGTTAATTGAGTTGATTTA	614				
QY	618	CAAGAAAGCTCGTCAGAAATATCACTTTCCCACTTGTGCTTGAGATTCAGATTCAGTGTG	677				
Db	615	TATGAAGCGCAAGCAAGCTGTTTACACTACAGTATTGGAGATTTCAGATGTTTGGG	674				
QY	678	CCTTGTGCTACTTTTCTTCTGCTGTCATTGTCATGTTGGAGATTTTAAGGTGAT	737				
Db	675	CATTCTGCTACTGTTTTTTTGGCACTATTGGAATGATTTCAACAAGGATTTCCAGGCGAT	734				
QY	738	AGCAAGAGAAGCAAGAGATTCAAGATTGGAGATTCAGTGTGTTTACTATGCAATGATAGT	797				
Db	735	TCCAGGGAAGCAAGTCAATATGCAATTGGAGAATCTAAG---TATTATATGTTACTAGT	791				
QY	798	GATCACAGGAATATATGGCAAGGTTTCTTCTTAGGAGCCATAGGAGATTCTGTTTCTGTC	857				
Db	792	ATGCTGTGCCAATTTTGGCAATTCGCGCTCTCGGGGTTAGTTGGAGTTATTTTTTATTC	851				
QY	858	ATCATCACTAGCTTCTGCTGTTCTGATAAGTGTCTCTCCGCTGACTGAAAGTTTTCGC	917				
Db	852	TTCTCTCTTACTCTCTGGAATATATAGCGGCTTTTCTTCTCTGTTTACGAAGTTTGGC	911				
QY	918	CGTCTGTTTGTTCGGGAGAAAGTTTCAGGAGAGAAAGGTTCTCTTACTCTTCTCTCT	977				
Db	912	TGTAATTTTGTTCACGAAAAA--TTTCAGCGCAAAAAGGAGTAGCTATTTTTTCTAGCTCT	969				
QY	978	TTGGGGAATTTGTCTCTTA	995				
Db	970	ATGGGGAATTTGTTTTCATA	987				
RESULT 14							
CK252005							
LOCUS							
DEFINITION							
ACCESSION							
VERSION							
KEYWORDS							
SOURCE							
ORGANISM							
REFERENCE							
AUTHORS							
TITLE							
JOURNAL							
COMMENT							
FEATURES							
source							
ORIGIN							
Query Match							
Best Local Similarity							
Matches							
QY	18	GAAGAAGATGAAGATGGTTTGAATATCAATAACTGTATTATCTTACCTCCTACATATAGGAACATG	77				
Db	39	GAATATGAAGAAGCTTATCTTCTTAATTTCAACATAATTATGTCTTCAATCGGCAATG	98				
QY	78	TGAGGTCTCTTGTGTTAACTCGTCTCTACCTTCAACATGCGGAAACGAATCTGGTTTCAT	137				
Db	99	CGTGTGCTTTTAACTCTCGACTCTATTTCATTATGCGGCGCAAGAAATTTGGATACC	158				
QY	138	GAGCTTCTTATCAACCGCTGGTTTTCATATFCTTCATCCCTCTCTTGGTCTCTCTTCT	197				
Db	159	CAGCGTATTACAACTGTTGGTTGTCCAAATTTCTCATCCCTCTAGCCATAGCCCTATT	218				
QY	198	CAGCCGTGCGCGGCAACCGCAACCTTAACACGCGGAAACAAACGAAGCT	257				
Db	219	CCAACCGCGGAAATTCAGG-----ACCGAGGCTTAAT	254				
QY	258	CTTCTCTCATGAAACTCTCTGTGGAGCGCTCCATTGTTCATAGGGTTGCTCACAGGACT	317				
Db	255	CGTCTTTATCACACGCGAAGATTTCATCGCATCTGCGGCGTGGGAATCATCGTGGTCT	314				
QY	318	TGACAACTACTTATATTTTACGGATTAGCATATCTGCCAGTTTCACTTCATCGTCTAT	377				
Db	315	CGATGGTTACTTAACTCATGGGACCCCGGAAATTAACCGTTTCACTTCGACTTAAT	374				



Db 852 TTCCTCTTTACTCTCTGGAATTATAGCGCTTTTTTATTCTCTGTTTACGAAAGTTTGGC 911  
Qy 918 CGTCGTTTGTTCGGGAGAGTTTTCAGGCAGACAGAAAGTGTCTCTCTACTTCTTCTCT 977  
Db 912 TGTAAATTT-TTTTCACGAAAATTTTCAAGCCGAAAAGAGTAGCTATTTTTTCTAGCTCT 970  
Qy 978 TTGGGGATT 986  
Db 971 ATGGGGATT 979

RESULT 15  
CK245030/c  
LOCUS  
DEFINITION CK245030 868 bp mRNA linear EST 12-DEC-2003  
EST728667 potato callus cDNA library, normalized and full-length  
Solanum tuberosum cDNA clone POCAB46 3' end, mRNA sequence.  
ACCESSION CK245030  
VERSION  
KEYWORDS  
SOURCE  
ORGANISM Solanum tuberosum (potato)  
Solanum tuberosum  
Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots;  
asterids; lamids; Solanales; Solanaceae; Solanum.  
REFERENCE Buell,C.R., Hart,A., Zismann,V., Karamycheva,S.A. and Baker,B.  
AUTHORS Generation of ESTs from potato callus tissue  
TITLE Unpublished (2003)  
JOURNAL  
COMMENT Other\_ESTs: EST728668 EST728669  
Contact: Robin Buell  
The Institute for Genomic Research  
9712 Medical Center Dr, Rockville, MD 20850, USA  
Email: potato-array@tigr.org  
Clones can be requested from TIGR via potato@tigr.org  
Seq primer: TTTTTTTTTTTTTTTTTTTTT.

FEATURES  
source  
1. .868  
/organism="Solanum tuberosum"  
/mol\_type="mRNA"  
/cultivar="Kennebec"  
/db\_xref="taxon:4113"  
/clone="POCAB46"  
/tissue\_type="callus"  
/lab\_host="DH10B-Tona"  
/clone\_lib="potato callus cDNA library, normalized and  
full-length"  
/notes="Vector: pCMVSPORT6.1; Site 1: EcoRI; Site 2: NotI;  
supplier: RNA was isolated from Solanum tuberosum var.  
Kennebec callus tissue grown on solid media."

ORIGIN  
Query Match 21.5%; Score 263; DB 14; Length 868;  
Best Local Similarity 60.6%; Pred. No. 7.5e-49;  
Matches 450; Conservative 0; Mismatches 290; Indels 3; Gaps 1;  
Qy 285 CGCTCCATTGTCATAGGGTTGCTCAGGACCTTGACAACTACTATATTCTTACGGATT 344  
Db 868 CGCATCTGCCGGCGTGGAAATCATCGTCGGTCTCGATGGTTACTTAACTCATGGGGACC 809  
Qy 345 AGCATATCTGCCAGTTTCAACTTTCATCGCTCATAATCGGAATCTCAACTAGCTTTCACGC 404  
Db 808 CGCGAAATTACCGTTTCACTTCGACTTAATCAACGGACCTCACTTCGGTTCACTGC 749  
Qy 405 TCTCTTCGCTTTCTTGTTAGTCAAGCAAAAGTTTCACTTCGTTTCCATAAACCGCGTGT 464  
Db 748 GCATTTTGCTGTGCTTATAGTTTAAACAGAAATTTGACAGCGCTATTTCGACGAATTCGGTGT 689  
Qy 465 TTTGTTGACGGTTGGTATCGGGATCTTGGGTTACACAGTGTGAGACAAACCGGCTAA 524  
Db 688 TTTTGCTTATCGCGGAGCTCGCATTTTAGCTCTCCGGCGGAACGGTGACCGCGGCGCG 629  
Qy 525 GGAGAGCAAGAAAGATGTCGTGTGGTTCTTTGATGACCTGTGCTTCACTCTCTCTCTA 584

Db 628 CGAGTCCACGAAGGATATATGTTGGGGTTTGTGATGACGTTTATCGGTGCGGTGTTATA 569  
QY 585 TGCTTTTATATATACCGCTCGTTGAGCTAACTTACAAGAAAGCTCGTCAAGAAATFCACTTT 644  
Db 568 TGGATTGATTTGCGTTAAATGAGTTGATTTATATGAAGCGGAAGCAAGCTGTACTTA 509  
QY 645 CCCACTTGTGCTTGAGATTCAGATGCTCATGTCGCTTGCCTGCTACTTTTTCGTGTCAT 704  
Db 508 CACTACAGTATTTGGAGATTCAGATGGTTTGGGCATTTCTGCTACTGTTTTTGCACAT 449  
QY 705 TGGCATGTTTCATCGTTGGAGATTTTAAAGTGATAGTACAGAGAAGCAAGAGAGTTCAAGAT 764  
Db 448 TGGAAATGATTCACAAGGATTTCCAGGCGATTTCCAGGCGATTTCCAGGGAAGCAAGTCAATATGCAAT 389  
QY 765 TGGAGGATCAGTGTGTTTTACTATGCATTCATGATGATCACAGGAATAATATGGCAAGGTTT 824  
Db 388 TGGAGAACTAAGT--ATTATATGCTACTAGTATGGTGTGCCATTTATTGGCAATTTCGC 332  
QY 825 CTTCTTAGGAGCCATAGGGAATGTTTTGTGCTATCATCACTAGCTTCTGTTCTGAT 884  
Db 331 GCTCCTGGGGTTAGTTGGAGTTAATTTTATTTCTTCTTCTTCTCTCTGGAATTTATAGG 272  
QY 885 AAGTCTTCTGCTTCGCTGACTGAAAGTTTTCGCCGTGCTTTGTTTCCGGGAGAAAGTTTCA 944  
Db 271 CGCTTTTCT 212  
QY 945 GGCAGAGAAAGGTCT 1004  
Db 211 AGCCGAAAGGAGTAGCTATTTTCTAGCTCTATGGGGATTTGTTTCATATATTTTATGG 152  
QY 1005 CGAGTTTAAATCCGGCAAGAAAG 1027  
Db 151 TGAATATGAACAAGGGAAGAGG 129

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM nucleic - nucleic search, using sw model

Run on: September 25, 2004, 10:08:56 ; Search time 549 Seconds  
(without alignments)  
9479.131 Million cell updates/sec

Title: US-09-913-767-1  
Perfect score: 1225  
Sequence: 1 aaacagcaagcagcaagaa.....aattgaaagatatgagct 1225

Scoring table: OLIGO NUC

Gapop 60.0 , Gapext 60.0

Searched: 3373863 seqs, 2124099041 residues

Word size : 0

Total number of hits satisfying chosen parameters: 6747726

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Listing first 1000 summaries

Database : N Geneseq\_29Jan04.\*

- 1: Geneseqn1980s.\*
- 2: Geneseqn1980s.\*
- 3: Geneseqn2000s.\*
- 4: Geneseqn2001as.\*
- 5: Geneseqn2001bs.\*
- 6: Geneseqn2002s.\*
- 7: Geneseqn2003as.\*
- 8: Geneseqn2003bs.\*
- 9: Geneseqn2003cs.\*
- 10: Geneseqn2004s.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1072	87.5	1225	3	AA97919 A. thalia
2	776	63.3	2586	3	AAC44184 Arabidops
3	253	20.7	493	3	AAC36831 Arabidops
4	21	1.7	729	7	ACA43388 Arabidops
C 5	21	1.7	2791	6	ABQ70805 Prokaryot
6	21	1.7	25392	4	AAK82159 Histaria
7	21	1.7	110000	6	ABA03041 Human imm
C 8	20	1.6	1017	7	ACA36531 Continuation (21 o
9	20	1.6	16488	6	AD25856 Prokaryot
10	20	1.6	16488	6	AD25856 Human dip
C 11	20	1.6	42738	4	AAK25895 Human dip
12	19	1.6	435	9	AD859081 A. thalia
13	19	1.6	524	6	ABQ38560 Oligonuc
C 14	19	1.6	524	6	ABQ38561 Oligonuc
15	19	1.6	644	3	AAC33529 Arabidops
16	19	1.6	738	4	AAH03331 Human CDN
C 17	19	1.6	813	3	AAC46918 Arabidops
18	19	1.6	915	5	ABV22998 Human pro
19	19	1.6	1319	6	ABZ16735 Arabidops
20	19	1.6	1319	7	ADA69110 Arabidops
21	19	1.6	1696	4	AAS03039 Human dia
C 22	19	1.6	2060	4	AAH13636 Human CDN
23	19	1.6	2132	4	AAH13691 Human CDN

19	1.6	2189	4	AAK51895	AAK51895 Human pol
19	1.6	2739	4	AAK52879	AAK52879 Human pol
25	19	3634	6	ABN95918	ABN95918 Gene #241
C 27	19	3647	2	AAV00085	AAV00085 Rat hexok
C 28	19	3647	2	AAV00159	AAV00159 Rat hexok
C 29	19	3653	6	AAI67596	AAI67596 Rat type
C 30	19	3653	7	ABT41828	ABT41828 Toxicity
C 31	19	3653	9	ADB58104	ADB58104 Toxicity
C 32	19	3653	9	ADB52585	ADB52585 Primary r
C 33	19	3653	9	ABL33359	ABL33359 Human imm
34	19	6419	6	ABL33359	ABL33359 DNA trans
35	19	6419	6	ABK28378	ABK28378 DNA trans
36	19	10205	6	ABK31274	ABK31274 Signal tr
37	19	10205	6	ABL70235	ABL70235 Chemical
38	19	31034	4	ABL13678	ABL13678 Drosophil
C 39	19	76798	6	ABN97454	ABN97454 Gene #395
C 40	18	295	5	ABV13479	ABV13479 Human pro
C 41	18	300	2	AAZ14679	AAZ14679 Human gen
C 42	18	304	3	AAA45946	AAA45946 Human met
C 43	18	346	2	AAV88758	AAV88758 EST clone
C 44	18	410	7	ABX49144	ABX49144 Bovine ES
C 45	18	452	6	ABL93879	ABL93879 Arabidops
C 46	18	471	6	ABV96522	ABV96522 Human pan
C 47	18	537	6	ABV88829	ABV88829 Human col
C 48	18	597	5	AAK63214	AAK63214 Human pur
C 49	18	606	3	AAK33596	AAK33596 Arabidops
C 50	18	619	4	AAK22848	AAK22848 Human CDN
C 51	18	631	3	AAF07955	AAF07955 Fusarium
52	18	664	4	AAK22612	AAK22612 Human CDN
53	18	687	4	AAK05556	AAK05556 Human CDN
54	18	701	5	AAH5831	AAH5831 Human SCN
55	18	798	3	AAK48613	AAK48613 Arabidops
56	18	807	3	AAK35092	AAK35092 Arabidops
57	18	821	4	AAH34736	AAH34736 Human col
58	18	914	2	AAK27423	AAK27423 Human sec
C 59	18	914	8	ADA07302	ADA07302 Human CDN
60	18	953	4	AAH78213	AAH78213 Nucleotid
61	18	1033	6	ABN98419	ABN98419 Arabidops
62	18	1236	2	AAK04444	AAK04444 Yeast imm
63	18	1380	7	ACC61183	ACC61183 Gene sequ
64	18	1434	4	AAK54345	AAK54345 Pseudomon
C 65	18	1434	7	ACA42721	ACA42721 Prokaryot
C 66	18	1443	6	ABA04916	ABA04916 Human RNA
C 67	18	1481	9	AAK06453	AAK06453 Arabidops
C 68	18	1481	9	ADE37166	ADE37166 Plant Vie
C 69	18	1670	3	AAK40003	AAK40003 Arabidops
C 70	18	1842	6	ABN59677	ABN59677 Novel hum
C 71	18	1884	4	AAK69143	AAK69143 Human imm
C 72	18	1967	4	AAK40960	AAK40960 CDNA enco
C 73	18	2000	6	ABZ17476	ABZ17476 Arabidops
74	18	2207	6	ABK63542	ABK63542 Rat sequ
75	18	2207	9	ADB59071	ADB59071 Toxicity
76	18	2207	9	ADB53155	ADB53155 Primary r
77	18	2234	4	AAH15955	AAH15955 Human CDN
C 78	18	2240	9	ABZ09864	ABZ09864 Human 5'
C 79	18	2240	9	ADE84018	ADE84018 5' regula
C 80	18	2323	4	AAK94598	AAK94598 Human ful
C 81	18	2358	4	ABL04982	ABL04982 Drosophil
C 82	18	2518	7	ABA00664	ABA00664 Human ENZ
C 83	18	2704	4	ABL13675	ABL13675 Drosophil
C 84	18	2723	4	ABL20387	ABL20387 Drosophil
C 85	18	2991	4	AAF03849	AAF03849 Novel hum
C 86	18	3131	4	AAH11170	AAH11170 Human mel
C 87	18	3231	4	ABL11375	ABL11375 Drosophil
C 88	18	3348	3	AAK50467	AAK50467 Arabidops
C 89	18	3365	4	AAH17203	AAH17203 Human mel
C 90	18	3372	6	ABA04908	ABA04908 Human RNA
C 91	18	3446	9	ADC30823	ADC30823 Human nov
C 92	18	3840	6	ABK99921	ABK99921 DNA encod
93	18	3967	4	ABL26340	ABL26340 Drosophil
94	18	4001	7	ADA20411	ADA20411 Prostata
95	18	4001	7	ADA84218	ADA84218 Human ren
96	18	4459	2	AAV99094	AAV99094 DNA methy

97	18	1.5	5562	6	ABL33934	Human imm	170	1.4	327	2	AAT68134	H. pylori
98	18	1.5	5562	6	ABL92294	Chemical	171	1.4	329	3	AAC13606	Human sec
99	18	1.5	5562	6	ABL49363	Human pol	172	1.4	339	3	AAC13606	Eucalyptu
100	18	1.5	5822	6	ABL33097	Human imm	173	1.4	351	6	ABQ79443	M. capsul
101	18	1.5	6120	4	ABQ46732	Tumour su	174	1.4	352	6	ABQ90100	Abq90100
102	18	1.5	6120	6	ABK31488	Signal tr	175	1.4	360	3	ABQ95800	Abq95800
103	18	1.5	6120	6	ABL70461	Chemical	176	1.4	360	3	AAA671134	Eucalyptu
104	18	1.5	6120	6	ABN80276	Human gen	177	1.4	360	3	AAA67185	E. grandi
105	18	1.5	6120	6	ABN80276	Human che	178	1.4	360	3	AAA67185	E. grandi
106	18	1.5	7111	6	ABL34210	Human imm	179	1.4	360	3	AAA67185	E. grandi
107	18	1.5	7384	6	ABL32747	Human imm	180	1.4	360	3	AAA67185	E. grandi
108	18	1.5	8229	3	AAH93450	Human APC	181	1.4	360	3	AAA67185	E. grandi
109	18	1.5	8532	2	AAV56447	Human APC	182	1.4	360	3	AAA67185	E. grandi
110	18	1.5	8532	3	AAH93449	Human APC	183	1.4	360	3	AAA67185	E. grandi
111	18	1.5	8532	9	ADC54132	Eukaryoti	184	1.4	360	3	AAA67185	E. grandi
112	18	1.5	8532	9	ADC54132	Eukaryoti	185	1.4	360	3	AAA67185	E. grandi
113	18	1.5	9606	2	AAQ27234	Adenomato	186	1.4	360	3	AAA67185	E. grandi
114	18	1.5	9606	2	AAQ27234	Adenomato	187	1.4	360	3	AAA67185	E. grandi
115	18	1.5	9606	2	AAQ27234	Adenomato	188	1.4	360	3	AAA67185	E. grandi
116	18	1.5	9606	2	AAQ27234	Adenomato	189	1.4	360	3	AAA67185	E. grandi
117	18	1.5	9606	2	AAQ27234	Adenomato	190	1.4	360	3	AAA67185	E. grandi
118	18	1.5	9606	2	AAQ27234	Adenomato	191	1.4	360	3	AAA67185	E. grandi
119	18	1.5	9606	2	AAQ27234	Adenomato	192	1.4	360	3	AAA67185	E. grandi
120	18	1.5	9606	2	AAQ27234	Adenomato	193	1.4	360	3	AAA67185	E. grandi
121	18	1.5	9606	2	AAQ27234	Adenomato	194	1.4	360	3	AAA67185	E. grandi
122	18	1.5	9606	2	AAQ27234	Adenomato	195	1.4	360	3	AAA67185	E. grandi
123	18	1.5	9606	2	AAQ27234	Adenomato	196	1.4	360	3	AAA67185	E. grandi
124	18	1.5	9606	2	AAQ27234	Adenomato	197	1.4	360	3	AAA67185	E. grandi
125	18	1.5	9606	2	AAQ27234	Adenomato	198	1.4	360	3	AAA67185	E. grandi
126	18	1.5	9606	2	AAQ27234	Adenomato	199	1.4	360	3	AAA67185	E. grandi
127	18	1.5	9606	2	AAQ27234	Adenomato	200	1.4	360	3	AAA67185	E. grandi
128	18	1.5	9606	2	AAQ27234	Adenomato	201	1.4	360	3	AAA67185	E. grandi
129	18	1.5	9606	2	AAQ27234	Adenomato	202	1.4	360	3	AAA67185	E. grandi
130	18	1.5	9606	2	AAQ27234	Adenomato	203	1.4	360	3	AAA67185	E. grandi
131	18	1.5	9606	2	AAQ27234	Adenomato	204	1.4	360	3	AAA67185	E. grandi
132	18	1.5	9606	2	AAQ27234	Adenomato	205	1.4	360	3	AAA67185	E. grandi
133	18	1.5	9606	2	AAQ27234	Adenomato	206	1.4	360	3	AAA67185	E. grandi
134	18	1.5	9606	2	AAQ27234	Adenomato	207	1.4	360	3	AAA67185	E. grandi
135	18	1.5	9606	2	AAQ27234	Adenomato	208	1.4	360	3	AAA67185	E. grandi
136	18	1.5	9606	2	AAQ27234	Adenomato	209	1.4	360	3	AAA67185	E. grandi
137	18	1.5	9606	2	AAQ27234	Adenomato	210	1.4	360	3	AAA67185	E. grandi
138	18	1.5	9606	2	AAQ27234	Adenomato	211	1.4	360	3	AAA67185	E. grandi
139	18	1.5	9606	2	AAQ27234	Adenomato	212	1.4	360	3	AAA67185	E. grandi
140	18	1.5	9606	2	AAQ27234	Adenomato	213	1.4	360	3	AAA67185	E. grandi
141	18	1.5	9606	2	AAQ27234	Adenomato	214	1.4	360	3	AAA67185	E. grandi
142	18	1.5	9606	2	AAQ27234	Adenomato	215	1.4	360	3	AAA67185	E. grandi
143	18	1.5	9606	2	AAQ27234	Adenomato	216	1.4	360	3	AAA67185	E. grandi
144	18	1.5	9606	2	AAQ27234	Adenomato	217	1.4	360	3	AAA67185	E. grandi
145	18	1.5	9606	2	AAQ27234	Adenomato	218	1.4	360	3	AAA67185	E. grandi
146	18	1.5	9606	2	AAQ27234	Adenomato	219	1.4	360	3	AAA67185	E. grandi
147	18	1.5	9606	2	AAQ27234	Adenomato	220	1.4	360	3	AAA67185	E. grandi
148	18	1.5	9606	2	AAQ27234	Adenomato	221	1.4	360	3	AAA67185	E. grandi
149	18	1.5	9606	2	AAQ27234	Adenomato	222	1.4	360	3	AAA67185	E. grandi
150	18	1.5	9606	2	AAQ27234	Adenomato	223	1.4	360	3	AAA67185	E. grandi
151	18	1.5	9606	2	AAQ27234	Adenomato	224	1.4	360	3	AAA67185	E. grandi
152	18	1.5	9606	2	AAQ27234	Adenomato	225	1.4	360	3	AAA67185	E. grandi
153	18	1.5	9606	2	AAQ27234	Adenomato	226	1.4	360	3	AAA67185	E. grandi
154	18	1.5	9606	2	AAQ27234	Adenomato	227	1.4	360	3	AAA67185	E. grandi
155	18	1.5	9606	2	AAQ27234	Adenomato	228	1.4	360	3	AAA67185	E. grandi
156	18	1.5	9606	2	AAQ27234	Adenomato	229	1.4	360	3	AAA67185	E. grandi
157	18	1.5	9606	2	AAQ27234	Adenomato	230	1.4	360	3	AAA67185	E. grandi
158	18	1.5	9606	2	AAQ27234	Adenomato	231	1.4	360	3	AAA67185	E. grandi
159	18	1.5	9606	2	AAQ27234	Adenomato	232	1.4	360	3	AAA67185	E. grandi
160	18	1.5	9606	2	AAQ27234	Adenomato	233	1.4	360	3	AAA67185	E. grandi
161	18	1.5	9606	2	AAQ27234	Adenomato	234	1.4	360	3	AAA67185	E. grandi
162	18	1.5	9606	2	AAQ27234	Adenomato	235	1.4	360	3	AAA67185	E. grandi
163	18	1.5	9606	2	AAQ27234	Adenomato	236	1.4	360	3	AAA67185	E. grandi
164	18	1.5	9606	2	AAQ27234	Adenomato	237	1.4	360	3	AAA67185	E. grandi
165	18	1.5	9606	2	AAQ27234	Adenomato	238	1.4	360	3	AAA67185	E. grandi
166	18	1.5	9606	2	AAQ27234	Adenomato	239	1.4	360	3	AAA67185	E. grandi
167	18	1.5	9606	2	AAQ27234	Adenomato	240	1.4	360	3	AAA67185	E. grandi
168	18	1.5	9606	2	AAQ27234	Adenomato	241	1.4	360	3	AAA67185	E. grandi
169	18	1.5	9606	2	AAQ27234	Adenomato	242	1.4	360	3	AAA67185	E. grandi



389	17	1.4	1835	4	AAI60297	Human pol	17	1.4	3518	5	AA02694	Human ade
C 390	17	1.4	1845	7	ADA69396	Rice gene	17	1.4	3561	4	AB128626	Drosophil
C 391	17	1.4	1868	4	AAK94482	Human ful	17	1.4	3567	2	AAK13164	Enterococ
C 392	17	1.4	1890	4	AAH16529	Human cdn	17	1.4	3567	6	ABS98959	Enterococ
C 393	17	1.4	1895	4	AAI58511	Human pol	17	1.4	3678	7	ADA13391	Human int
C 394	17	1.4	1895	8	ADB48487	Novel hum	17	1.4	3678	7	ADL30242	Drosophil
C 395	17	1.4	1917	3	AA424272	Novel hum	17	1.4	3767	4	AB115570	Drosophil
C 396	17	1.4	1933	1	AAAN90120	Human lip	17	1.4	3767	6	AB115570	Drosophil
C 397	17	1.4	1938	4	AB112035	Drosophil	17	1.4	3769	6	AD28057	Human ade
C 398	17	1.4	1953	3	AA97449	Chicory g	17	1.4	3824	4	AB118246	Drosophil
C 399	17	1.4	1954	2	AAAT13382	Pea ADP-	17	1.4	3854	5	ABV28651	Human pro
C 400	17	1.4	1954	2	AAV35052	Pea ADP-	17	1.4	3854	5	ABV28651	Human pro
C 401	17	1.4	1985	4	ABU10799	Drosophil	17	1.4	3855	5	ABV21532	Human pro
C 402	17	1.4	1991	2	AAQ48764	GAT-3 tra	17	1.4	3855	5	ABV21532	Human pro
C 403	17	1.4	1991	2	AAQ48764	GAT-3 tra	17	1.4	3855	5	ABV21532	Human pro
C 404	17	1.4	2002	9	AAAF85439	Brono RNA	17	1.4	3855	5	ABV21532	Human pro
C 405	17	1.4	2022	7	ACSA54012	Prokaryot	17	1.4	3855	5	ABV21532	Human pro
C 406	17	1.4	2051	5	AA574674	Human cod	17	1.4	3855	5	ABV21532	Human pro
C 407	17	1.4	2094	7	ADA52495	Human cod	17	1.4	3855	5	ABV21532	Human pro
C 408	17	1.4	2096	4	AAK94485	Human ful	17	1.4	3855	5	ABV21532	Human pro
C 409	17	1.4	2126	2	AAQ411229	Clone GP1	17	1.4	3855	5	ABV21532	Human pro
C 410	17	1.4	2160	4	AAH15898	Human cdn	17	1.4	3855	5	ABV21532	Human pro
C 411	17	1.4	2187	4	AAAF81507	White lup	17	1.4	3855	5	ABV21532	Human pro
C 412	17	1.4	2232	5	AA592576	DNA encod	17	1.4	3855	5	ABV21532	Human pro
C 413	17	1.4	2232	5	AA583344	DNA encod	17	1.4	3855	5	ABV21532	Human pro
C 414	17	1.4	2232	5	AA583344	DNA encod	17	1.4	3855	5	ABV21532	Human pro
C 415	17	1.4	2262	3	AAV06954	Alcaligen	17	1.4	3855	5	ABV21532	Human pro
C 416	17	1.4	2262	3	AA555491	Human TRA	17	1.4	3855	5	ABV21532	Human pro
C 417	17	1.4	2262	5	AA001949	Human TNF	17	1.4	3855	5	ABV21532	Human pro
C 418	17	1.4	2262	6	AB165876	Lung canc	17	1.4	3855	5	ABV21532	Human pro
C 419	17	1.4	2262	6	AB165876	Lung canc	17	1.4	3855	5	ABV21532	Human pro
C 420	17	1.4	2262	6	AB165876	Lung canc	17	1.4	3855	5	ABV21532	Human pro
C 421	17	1.4	2262	7	ACD13346	DNA encod	17	1.4	3855	5	ABV21532	Human pro
C 422	17	1.4	2262	8	ADB80997	RING-SH c	17	1.4	3855	5	ABV21532	Human pro
C 423	17	1.4	2374	3	AA529280	Arabidops	17	1.4	3855	5	ABV21532	Human pro
C 424	17	1.4	2420	7	AD567778	Human cod	17	1.4	3855	5	ABV21532	Human pro
C 425	17	1.4	2436	7	AAH53483	Human cod	17	1.4	3855	5	ABV21532	Human pro
C 426	17	1.4	2489	4	AAH17700	Human cdn	17	1.4	3855	5	ABV21532	Human pro
C 427	17	1.4	2497	6	AB583359	CDNA enco	17	1.4	3855	5	ABV21532	Human pro
C 428	17	1.4	2561	6	AA562313	CDNA sequ	17	1.4	3855	5	ABV21532	Human pro
C 429	17	1.4	2592	5	AA575052	DNA encod	17	1.4	3855	5	ABV21532	Human pro
C 430	17	1.4	2598	7	AD568117	Arabidops	17	1.4	3855	5	ABV21532	Human pro
C 431	17	1.4	2634	6	AB211902	Human pol	17	1.4	3855	5	ABV21532	Human pro
C 432	17	1.4	2650	6	AA561689	Lung smal	17	1.4	3855	5	ABV21532	Human pro
C 433	17	1.4	2650	3	AAZ93150	Sequence	17	1.4	3855	5	ABV21532	Human pro
C 434	17	1.4	2651	9	ADB62830	Human cdn	17	1.4	3855	5	ABV21532	Human pro
C 435	17	1.4	2667	7	ACA48411	Prokaryot	17	1.4	3855	5	ABV21532	Human pro
C 436	17	1.4	2676	4	AB111753	Drosophil	17	1.4	3855	5	ABV21532	Human pro
C 437	17	1.4	2676	7	ABT43044	Human neu	17	1.4	3855	5	ABV21532	Human pro
C 438	17	1.4	2694	6	AAH22131	Arabidops	17	1.4	3855	5	ABV21532	Human pro
C 439	17	1.4	2738	4	AAH14243	Human cdn	17	1.4	3855	5	ABV21532	Human pro
C 440	17	1.4	2770	3	AAZ93410	Human sec	17	1.4	3855	5	ABV21532	Human pro
C 441	17	1.4	2808	4	AB114438	Drosophil	17	1.4	3855	5	ABV21532	Human pro
C 442	17	1.4	2880	3	AAA49189	CDNA enco	17	1.4	3855	5	ABV21532	Human pro
C 443	17	1.4	2892	4	AAH17602	Human cdn	17	1.4	3855	5	ABV21532	Human pro
C 444	17	1.4	2968	3	AA588853	Human dif	17	1.4	3855	5	ABV21532	Human pro
C 445	17	1.4	2968	3	AAA98134	Human dif	17	1.4	3855	5	ABV21532	Human pro
C 446	17	1.4	2968	6	AAAD29964	Human pha	17	1.4	3855	5	ABV21532	Human pro
C 447	17	1.4	3079	4	AB121484	Drosophil	17	1.4	3855	5	ABV21532	Human pro
C 448	17	1.4	3084	9	ADB62978	Human cdn	17	1.4	3855	5	ABV21532	Human pro
C 449	17	1.4	3134	6	ABQ54736	Human ova	17	1.4	3855	5	ABV21532	Human pro
C 450	17	1.4	3159	6	ABA91457	Human Dow	17	1.4	3855	5	ABV21532	Human pro
C 451	17	1.4	3159	9	ADE37338	Human MCI	17	1.4	3855	5	ABV21532	Human pro
C 452	17	1.4	3184	6	AAAD30157	Colon ade	17	1.4	3855	5	ABV21532	Human pro
C 453	17	1.4	3184	6	AB161768	Lung canc	17	1.4	3855	5	ABV21532	Human pro
C 454	17	1.4	3184	6	AB165187	Lung canc	17	1.4	3855	5	ABV21532	Human pro
C 455	17	1.4	3186	7	ACD13348	Human DNA	17	1.4	3855	5	ABV21532	Human pro
C 456	17	1.4	3248	4	AB126846	Drosophil	17	1.4	3855	5	ABV21532	Human pro
C 457	17	1.4	3261	9	ADB62092	Human cdn	17	1.4	3855	5	ABV21532	Human pro
C 458	17	1.4	3312	6	ABN85298	Human ade	17	1.4	3855	5	ABV21532	Human pro
C 459	17	1.4	3320	7	ADA19396	Human ins	17	1.4	3855	5	ABV21532	Human pro
C 460	17	1.4	3321	9	ADC39123	Novel hum	17	1.4	3855	5	ABV21532	Human pro
C 461	17	1.4	3357	3	AAA53921	Type IV a	17	1.4	3855	5	ABV21532	Human pro







C 827	16	1.3	430	6	ABV97075	Human pan	900	16	1.3	493	8	ACH19120	Human adu
C 828	16	1.3	431	8	ACH30272	Human tes	C 901	16	1.3	494	7	ACD95488	Human col
C 829	16	1.3	432	7	ACA31905	Prokaryot	C 902	16	1.3	495	7	AAC54171	Arabidops
C 830	16	1.3	432	7	ACA35484	Prokaryot	C 903	16	1.3	495	7	ABT41048	Toxicity
C 831	16	1.3	434	3	AAC07590	Human sec	C 904	16	1.3	495	7	ACA14073	Prokaryot
C 832	16	1.3	435	4	AAS56205	Salmonell	C 905	16	1.3	496	8	ACH16576	Human adu
C 833	16	1.3	435	7	ACA51197	Prokaryot	C 906	16	1.3	496	8	ACH26864	Human adu
C 834	16	1.3	437	4	AAL04214	Human rep	C 907	16	1.3	499	4	AAL22930	Human bre
C 835	16	1.3	438	5	ABV12187	Human pro	C 908	16	1.3	500	3	AAC95156	Cat flea
C 836	16	1.3	443	1	AAN70773	Sequence	C 909	16	1.3	501	9	ADD49434	Human lun
C 837	16	1.3	445	5	ABV42255	Human pro	C 910	16	1.3	501	9	ABX61412	Arabidops
C 838	16	1.3	445	5	ABV33332	Human pro	C 911	16	1.3	503	7	ADD16954	Arabidops
C 839	16	1.3	446	2	AAQ44247	HEJR-1 ge	C 912	16	1.3	503	4	AAI17472	DNA (Seq)
C 840	16	1.3	446	8	ACH31938	Human end	C 913	16	1.3	505	4	ABA62403	Probe #74
C 841	16	1.3	448	4	ABA08505	Human nov	C 914	16	1.3	505	4	ABA62403	Human foe
C 842	16	1.3	449	5	ABV54441	Human pro	C 915	16	1.3	505	4	Aai42381	Probe #11
C 843	16	1.3	453	3	AAC01867	Human sec	C 916	16	1.3	505	4	ABA29737	Probe #82
C 844	16	1.3	456	3	AAA43753	Human sec	C 917	16	1.3	505	4	AAC36621	Human bon
C 845	16	1.3	456	4	AAL20021	Human bre	C 918	16	1.3	505	4	AAC36621	Human bon
C 846	16	1.3	456	5	AAS80548	DNA encod	C 919	16	1.3	506	6	ABK44819	Human liv
C 847	16	1.3	457	6	ABN16656	Human ORF	C 920	16	1.3	506	9	ABE81515	cdna enco
C 848	16	1.3	458	4	AAL10600	Human bre	C 921	16	1.3	509	5	ABV49121	Arabidops
C 849	16	1.3	461	8	ACH41648	Human foe	C 922	16	1.3	513	9	ADC92075	Human pro
C 850	16	1.3	462	4	AAL92460	Human pol	C 923	16	1.3	513	7	ADC92075	E. faeciu
C 851	16	1.3	462	5	AAS88431	DNA encod	C 924	16	1.3	516	7	ACD96028	Human col
C 852	16	1.3	463	8	ACH47935	Human lun	C 925	16	1.3	517	6	ABK45093	cdna enco
C 853	16	1.3	464	5	AAS67747	DNA encod	C 926	16	1.3	521	3	ABN63297	Human can
C 854	16	1.3	464	8	ACH31954	Human end	C 927	16	1.3	523	3	AAC56975	Pinus rad
C 855	16	1.3	467	4	AAH05744	Human cdn	C 928	16	1.3	527	3	AAC04030	Human sec
C 856	16	1.3	471	4	AAI15295	Probe #52	C 929	16	1.3	527	4	AAH06154	Human cdn
C 857	16	1.3	471	4	ABA57077	Human foe	C 930	16	1.3	528	6	ABZ14402	Arabidops
C 858	16	1.3	471	4	AAI81718	Human pol	C 931	16	1.3	529	6	ABQ36362	Oligonucl
C 859	16	1.3	473	5	ABV03018	Human pro	C 932	16	1.3	531	3	ABQ36363	Arabidops
C 860	16	1.3	473	7	ACA14572	Prokaryot	C 933	16	1.3	531	3	AAC41248	Human bre
C 861	16	1.3	473	8	ACH47140	Human inf	C 934	16	1.3	535	5	AAI18979	Human bre
C 862	16	1.3	474	2	AAV87079	EST clone	C 935	16	1.3	535	4	AAK68286	Human G-P
C 863	16	1.3	474	5	ABV34079	Human pro	C 936	16	1.3	536	4	AAK68294	Human imm
C 864	16	1.3	474	5	ABV42944	Human pro	C 937	16	1.3	536	4	AAK68293	Human imm
C 865	16	1.3	474	5	ABV39324	Human pro	C 938	16	1.3	536	4	AAK68288	Human imm
C 866	16	1.3	477	8	ADA48069	Rice gene	C 939	16	1.3	536	4	AAK68290	Human imm
C 867	16	1.3	478	4	AAI10441	Human bre	C 940	16	1.3	536	4	AAK68286	Human imm
C 868	16	1.3	478	8	ACH26046	Human adu	C 941	16	1.3	536	4	AAK68295	Human imm
C 869	16	1.3	479	2	AAV87283	EST clone	C 942	16	1.3	536	4	AAK68289	Human imm
C 870	16	1.3	481	4	AAK61813	Human imm	C 943	16	1.3	536	4	AAK68291	Human imm
C 871	16	1.3	482	7	ACA14811	Prokaryot	C 944	16	1.3	536	4	AAK68287	Human imm
C 872	16	1.3	482	7	ACA14706	Prokaryot	C 945	16	1.3	537	4	ABA60956	Human foe
C 873	16	1.3	484	4	AAI12034	Probe #19	C 946	16	1.3	537	4	AAI40853	Probe #95
C 874	16	1.3	484	4	ABA53736	Human foe	C 947	16	1.3	537	4	AAK35136	Human bon
C 875	16	1.3	484	4	AAI33368	Probe #20	C 948	16	1.3	537	4	AAK09246	Human bra
C 876	16	1.3	484	4	AAK81065	Human imm	C 949	16	1.3	537	4	ABS34887	Human liv
C 877	16	1.3	484	4	AAK81066	Human imm	C 950	16	1.3	537	6	ABS09585	Human gen
C 878	16	1.3	484	4	AAK81064	Human imm	C 951	16	1.3	541	4	AAK56553	Human imm
C 879	16	1.3	484	4	ABA43304	Human bre	C 952	16	1.3	543	3	AAA66044	E. coli p
C 880	16	1.3	484	4	ABA33487	Probe #19	C 953	16	1.3	543	4	AAS24444	E. coli D
C 881	16	1.3	484	4	AAK27457	Human bon	C 954	16	1.3	543	4	AH84542	E. coli g
C 882	16	1.3	484	4	AAK02000	Human bra	C 955	16	1.3	543	7	ACA51516	Prokaryot
C 883	16	1.3	484	4	ABS27024	Human liv	C 956	16	1.3	543	7	ACA18586	Prokaryot
C 884	16	1.3	484	5	AAI01961	Probe #19	C 957	16	1.3	544	7	ACA48692	Prokaryot
C 885	16	1.3	484	5	ABV07751	Human pro	C 958	16	1.3	547	4	ABA61983	Human foe
C 886	16	1.3	484	6	ABS01950	Human gen	C 959	16	1.3	547	4	AAI41912	Probe #10
C 887	16	1.3	485	6	ABQ57813	Human col	C 960	16	1.3	547	4	ABA29473	Probe #79
C 888	16	1.3	486	7	ABX61025	Arabidops	C 961	16	1.3	547	4	AAC36196	Human bon
C 889	16	1.3	488	8	ACH15284	Human adu	C 962	16	1.3	547	4	AAK10294	Human bra
C 890	16	1.3	490	4	AAK65933	Human imm	C 963	16	1.3	547	4	ABS35885	Human liv
C 891	16	1.3	490	4	AAL03581	Human rep	C 964	16	1.3	547	6	ABS10304	Human gen
C 892	16	1.3	490	4	AAL03580	Human rep	C 965	16	1.3	548	2	AAI13604	Enterococ
C 893	16	1.3	490	4	AAL03579	Human rep	C 966	16	1.3	548	3	ABS99399	Enterococ
C 894	16	1.3	490	4	ABA07790	Human ova	C 967	16	1.3	549	3	AAC95221	Cat flea
C 895	16	1.3	490	4	ABA07789	Human ova	C 968	16	1.3	550	5	ABV21951	Human pro
C 896	16	1.3	490	4	ABA07788	Human ova	C 969	16	1.3	550	5	ABV27783	Human pro
C 897	16	1.3	490	4	AAH08039	Human cdn	C 970	16	1.3	550	5	ABV21519	Human pro
C 898	16	1.3	490	9	ADC55555	Drosophil	C 971	16	1.3	550	5	ABV27338	Human pro
C 899	16	1.3	492	7	ACA13704	Prokaryot	C 972	16	1.3	552	8	ACH22853	Human adu

## ALIGNMENTS

XX  
ps  
Claim 2. Page 11: 24pp: German.

```
QY 781 TACTATGCATGATGATGATCACAGGAATATATATGGCAAGGTTCTTCTTAGGAGCCATA 840
Db 781 TACTATGCATGATGATGATCACAGGAATATATGGCAAGGTTCTTCTTAGGAGCCATA 840
QY 841 GGGATTGTTGTTGTCATCATCACTAGCTTCTGGTGTCTGATAAGTGTTCGTTCCG 900
Db 841 GGGATTGTTGTTGTCATCATCACTAGCTTCTGGTGTCTGATAAGTGTTCGTTCCG 900
QY 901 GTGACTGAAGTTTTCGGCGTGGTTGTTTCGGGAGAAAGTTTCAGGCAGAAAGGTGTC 960
Db 901 GTGACTGAAGTTTTCGGCGTGGTTGTTTCGGGAGAAAGTTTCAGGCAGAAAGGTGTC 960
QY 961 TCTCTACTCTTCTCTTCTGGGATTTGTCCTTACTTCTACGCGAGTTTAAATCCGGC 1020
Db 961 TCTCTACTCTTCTCTTCTGGGATTTGTCCTTACTTCTACGCGAGTTTAAATCCGGC 1020
QY 1021 AAGAAAGTCTTGATAACTCAACCCCGGAGACAGAACTGCCTATTCTTCCAGTTAGT 1080
Db 1021 AAGAAAGTCTTGATAACTCAACCCCGGAGACAGAACTGCCTATTCTTCCAGTTAGT 1080
QY 1081 GATTATGTTCTTAATTTCTATTAACCTCTATACGATTATAACAGACGATTACTGTTATGTT 1140
Db 1081 GATTATGTTCTTAATTTCTATTAACCTCTATACGATTATAACAGACGATTACTGTTATGTT 1140
QY 1141 TTGTTCTCTAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
Db 1141 TTGTTCTCTAAATATATGATGATGATGATGATGATGATGATGATGATGATGATGATG 1200
QY 1201 AATAAAATTTGAAAGATATTGAGCT 1225
Db 1201 AATAAAATTTGAAAGATATTGAGCT 1225

RESULT 2
AAC44184
ID AAC44184 standard; DNA; 2586 BP.
XX AC AAC44184;
XX DT 18-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 41938.
XX KW Hybridisation assay; genetic mapping; gene expression control;
XX KW protein identification; signal transduction pathway; metabolic pathway;
XX KW promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 28-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 30-APR-1999; 99US-0132407P.
PR 04-MAY-1999; 99US-0132484P.
PR 05-MAY-1999; 99US-0132485P.
PR 06-MAY-1999; 99US-0132486P.
PR 07-MAY-1999; 99US-0132487P.
PR 11-MAY-1999; 99US-0132863P.
PR 14-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 18-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 24-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 28-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
```



```
OY 1065 TATCTTCAGTTAGTCATTATGTTGCTTAATTTCTATACTCTATA 1111
Db 2429 TATCTTCAGTTAGTCATTATGTTGCTTAATTTCTATACTCTATA 2475

RESULT 3
AAC36831
ID AAC36831 standard; DNA; 493 BP.
XX AC AAC36831;
XX DT 17-OCT-2000 (first entry)
XX DE Arabidopsis thaliana DNA fragment SEQ ID NO: 15222.
XX KW Hybridisation assay; genetic mapping; gene expression control;
KW protein identification; signal transduction pathway; metabolic pathway;
KW promoter; termination sequence; ss.
XX OS Arabidopsis thaliana.
XX PN EP1033405-A2.
XX PD 06-SEP-2000.
XX PF 25-FEB-2000; 2000EP-00301439.
XX 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 23-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
PR 16-APR-1999; 99US-0129845P.
PR 19-APR-1999; 99US-0130077P.
PR 21-APR-1999; 99US-0130449P.
PR 23-APR-1999; 99US-0130510P.
PR 28-APR-1999; 99US-0130891P.
PR 30-APR-1999; 99US-0131449P.
PR 30-APR-1999; 99US-0132048P.
PR 04-MAY-1999; 99US-0132240P.
PR 05-MAY-1999; 99US-0132484P.
PR 06-MAY-1999; 99US-0132486P.
PR 06-MAY-1999; 99US-0132487P.
PR 07-MAY-1999; 99US-0132863P.
PR 11-MAY-1999; 99US-0134256P.
PR 14-MAY-1999; 99US-0134218P.
PR 14-MAY-1999; 99US-0134219P.
PR 14-MAY-1999; 99US-0134221P.
PR 14-MAY-1999; 99US-0134370P.
PR 18-MAY-1999; 99US-0134768P.
PR 19-MAY-1999; 99US-0134941P.
PR 20-MAY-1999; 99US-0135124P.
PR 21-MAY-1999; 99US-0135353P.
PR 24-MAY-1999; 99US-0135629P.
PR 25-MAY-1999; 99US-0136021P.
PR 27-MAY-1999; 99US-0136392P.
PR 28-MAY-1999; 99US-0136782P.
PR 01-JUN-1999; 99US-0137222P.
PR 03-JUN-1999; 99US-0137528P.
PR 04-JUN-1999; 99US-0137502P.
PR 07-JUN-1999; 99US-0137724P.
PR 08-JUN-1999; 99US-0138094P.
PR 10-JUN-1999; 99US-0138540P.
PR 10-JUN-1999; 99US-0138847P.
PR 14-JUN-1999; 99US-0139119P.
PR 16-JUN-1999; 99US-0139452P.
PR 16-JUN-1999; 99US-0139453P.
PR 17-JUN-1999; 99US-0139492P.
PR 18-JUN-1999; 99US-0139454P.
PR 18-JUN-1999; 99US-0139455P.
PR 18-JUN-1999; 99US-0139456P.
PR 18-JUN-1999; 99US-0139457P.
PR 18-JUN-1999; 99US-0139458P.
PR 18-JUN-1999; 99US-0139459P.
PR 18-JUN-1999; 99US-0139460P.
PR 18-JUN-1999; 99US-0139461P.
PR 18-JUN-1999; 99US-0139462P.
PR 18-JUN-1999; 99US-0139463P.
PR 18-JUN-1999; 99US-0139750P.
PR 18-JUN-1999; 99US-0139763P.
PR 21-JUN-1999; 99US-0139817P.
PR 22-JUN-1999; 99US-0139899P.
PR 23-JUN-1999; 99US-0140353P.
PR 23-JUN-1999; 99US-0140354P.
PR 24-JUN-1999; 99US-0140695P.
PR 28-JUN-1999; 99US-0140823P.
PR 29-JUN-1999; 99US-0140991P.
PR 30-JUN-1999; 99US-0141287P.
PR 01-JUL-1999; 99US-0141842P.
PR 01-JUL-1999; 99US-0142154P.
PR 02-JUL-1999; 99US-0142055P.
PR 06-JUL-1999; 99US-0142390P.
PR 08-JUL-1999; 99US-0142803P.
PR 09-JUL-1999; 99US-0142920P.
PR 12-JUL-1999; 99US-0142977P.
PR 13-JUL-1999; 99US-0143542P.
PR 14-JUL-1999; 99US-0143624P.
PR 15-JUL-1999; 99US-0144005P.
PR 16-JUL-1999; 99US-0144085P.
PR 16-JUL-1999; 99US-0144086P.
PR 19-JUL-1999; 99US-0144325P.
PR 19-JUL-1999; 99US-0144331P.
PR 19-JUL-1999; 99US-0144332P.
PR 19-JUL-1999; 99US-0144333P.
PR 19-JUL-1999; 99US-0144334P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144352P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145085P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146389P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
```

QY	121	AAACGAATCTGGTTTCATGAGCTTCCTATCAACCGCTGGTTTTCCCAATCATCTCTCATCCCT	180
Db	157	AAACGAATCTGGTTTCATGAGCTTCCTATCAACCGCTGGTTTTCCCAATCATCTCTCATCCCT	216
QY	181	CTCTTGGTCTCTCTTCCTCAGCGCTCGCCGGGCAACCGCAACCTTAAACAACGGGAAAC	240
Db	217	CTCTTGGTCTCTCTTCCTCAGCGCTCGCCGGGCAACCGCAACCTTAAACAACGGGAAAC	276
QY	241	AGCGGAAAACAAGCTCTTCTCATGGAATCTCTCTGTGGAGCGCTCCATTGTGCTATA	300
Db	277	AGCGGAAAACAAGCTCTTCTCATGGAATCTCTCTGTGGAGCGCTCCATTGTGCTATA	336
QY	301	GGGTGCTCACAGGACTTGACAACTATTATATCTTACGGATTAGCATATCTGCCAGTT	360
Db	337	GGGTGCTCACAGGACTTGACAACTATTATATCTTACGGATTAGCATATCTGCCAGTT	396
QY	361	TCAACTTCATCGCTCATATCGGAATCTCAACTAGCTTTCAACGCTCTCTTCGCTTTCTTG	420
Db	397	TCAACTTCATCGCTCATATCGGAATCTCAACTAGCTTTCAACGCTCTCTTCGCTTTCTTG	456
QY	421	TTAGTCAAGCAAAAGTTCACTCCGTTCTCCATAAACG	457
Db	457	TTAGTCAAGCAAAAGTTCACTCCGTTCTCCATAAACG	493
RESULT 4			
ACA43388			
ID	ACA43388 standard; DNA; 729 BP.		
XX	ACA43388;		
AC	ACA43388;		
XX	19-JUN-2003 (first entry)		
DT	Prokaryotic essential gene #25045.		
XX	Antisense; ds; prokaryotic essential gene; cell proliferation;		
DE	drug design; gene.		
KW	Pasteurella multocida.		
XX	WO200277183-A2.		
OS	03-OCT-2002.		
XX	21-MAR-2002; 2002WO-US009107.		
PN	21-MAR-2001; 2001US-00815242.		
XX	06-SEP-2001; 2001US-00948993.		
PR	25-OCT-2001; 2001US-0342923P.		
PR	08-FEB-2002; 2002US-00072851.		
PR	06-MAR-2002; 2002US-0362699P.		
XX	(ELIT-) ELITRA PHARM INC.		
PA	Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen XL, Zyskind JW;		
XX	Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;		
PI	WPI; 2003-029926/02.		
XX	P-PSDB; ABU39518.		
DR	New antisense nucleic acids, useful for identifying proteins or screening		
DR	for homologous nucleic acids required for cellular proliferation to		
PT	isolate candidate molecules for rational drug discovery programs.		
PT	Claim 14; SEQ ID NO 31258; 1766pp; English.		
XX	The invention relates to an isolated nucleic acid comprising any one of		
PS	the 6213 antisense sequences given in the specification where expression		
CC	of the nucleic acid inhibits proliferation of a cell. Also included are:		
CC	(1) a vector comprising a promoter operably linked to the nucleic acid		
CC	encoding a polypeptide whose expression is inhibited by the antisense		
CC	nucleic acid; (2) a host cell containing the vector; (3) an isolated		

Qy	Seq	Score	DB 3	Length	493
16-AUG-1999	R	99US-0149368P.			
17-AUG-1999	R	99US-01494175P.			
18-AUG-1999	R	99US-0149426P.			
20-AUG-1999	R	99US-0149722P.			
20-AUG-1999	R	99US-0149723P.			
20-AUG-1999	R	99US-0149929P.			
20-AUG-1999	R	99US-0149902P.			
23-AUG-1999	R	99US-0149930P.			
23-AUG-1999	R	99US-0150566P.			
25-AUG-1999	R	99US-0150884P.			
26-AUG-1999	R	99US-0151065P.			
27-AUG-1999	R	99US-0151066P.			
27-AUG-1999	R	99US-0151080P.			
30-AUG-1999	R	99US-0151303P.			
31-AUG-1999	R	99US-0151438P.			
01-SEP-1999	R	99US-0151930P.			
07-SEP-1999	R	99US-0152363P.			
10-SEP-1999	R	99US-0153070P.			
13-SEP-1999	R	99US-0153758P.			
15-SEP-1999	R	99US-0154018P.			
16-SEP-1999	R	99US-0154039P.			
20-SEP-1999	R	99US-0154779P.			
22-SEP-1999	R	99US-0155139P.			
23-SEP-1999	R	99US-0155486P.			
24-SEP-1999	R	99US-0155659P.			
28-SEP-1999	R	99US-0156458P.			
29-SEP-1999	R	99US-0156596P.			
04-OCT-1999	R	99US-0157117P.			
05-OCT-1999	R	99US-0157753P.			
06-OCT-1999	R	99US-0157865P.			
07-OCT-1999	R	99US-0158029P.			
08-OCT-1999	R	99US-0158232P.			
12-OCT-1999	R	99US-0158369P.			
13-OCT-1999	R	99US-0159293P.			
13-OCT-1999	R	99US-0159294P.			
13-OCT-1999	R	99US-0159295P.			
14-OCT-1999	R	99US-0159329P.			
14-OCT-1999	R	99US-0159330P.			
14-OCT-1999	R	99US-0159331P.			
14-OCT-1999	R	99US-0159637P.			
14-OCT-1999	R	99US-0159638P.			
18-OCT-1999	R	99US-0159584P.			
21-OCT-1999	R	99US-0160741P.			
21-OCT-1999	R	99US-0160767P.			
21-OCT-1999	R	99US-0160768P.			
21-OCT-1999	R	99US-0160770P.			
21-OCT-1999	R	99US-0160814P.			
21-OCT-1999	R	99US-0160815P.			
22-OCT-1999	R	99US-0160980P.			
22-OCT-1999	R	99US-0160981P.			
22-OCT-1999	R	99US-0160989P.			
25-OCT-1999	R	99US-0161404P.			
25-OCT-1999	R	99US-0161405P.			
25-OCT-1999	R	99US-0161406P.			
26-OCT-1999	R	99US-0161359P.			
26-OCT-1999	R	99US-0161360P.			
26-OCT-1999	R	99US-0161361P.			
28-OCT-1999	R	99US-0161920P.			
28-OCT-1999	R	99US-0161992P.			
28-OCT-1999	R	99US-0161993P.			
29-OCT-1999	R	99US-0162142P.			

Query Match 20.7%; Score 253; DB 3; Length 493;

Best Local Similarity 99.1%; Pred. No. 3.7e-115; Indels 0; Gaps 0;

Matches 453; Conservative 0; Mismatches 4;

Qy	Seq	Score	DB 3	Length	493
1	AAAAAGCAAGCAGCAAGCAAGATGAAGATGGTTTGATATCATATAAATCTGTTATTATC	60			
37	AAAAAGCAAGCAGCAAGCAAGATGAAGATGGTTTGATATCATATAAATCTGTTATTATC	96			

CC polypeptide or its fragment whose expression is inhibited by the  
 CC antiseptic nucleic acid; (4) an antibody capable of specifically binding  
 CC the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular  
 CC proliferation or the activity of a gene in an operon required for  
 CC proliferation; (7) identifying a compound that influences the activity of  
 CC the gene product or that has an activity against a biological pathway  
 CC required for proliferation, or that inhibits cellular proliferation; (8)  
 CC identifying a gene required for cellular proliferation or the biological  
 CC pathway in which a proliferation-required gene or its gene product lies  
 CC or a gene on which the test compound that inhibits proliferation of an  
 CC organism acts; (9) manufacturing an antibiotic; (10) profiling a  
 CC compound's activity; (11) a culture comprising strains in which the gene  
 CC product is overexpressed or underexpressed; (12) determining the extent  
 CC to which each of the strains is present in a culture or collection of  
 CC strains; or (13) identifying the target of a compound that inhibits the  
 CC proliferation of an organism. The antiseptic nucleic acids are useful for  
 CC identifying proteins or screening for homologous nucleic acids required  
 CC for cellular proliferation to isolate candidate molecules for rational  
 CC drug discovery programs, or for screening homologous nucleic acids  
 CC required for proliferation in cells other than *S. aureus*, *S. typhimurium*,  
 CC *K. pneumoniae* or *P. aeruginosa*. The present sequence is one of the target  
 CC prokaryotic essential genes. Note: The sequence data for this patent did  
 CC not form part of the printed specification, but was obtained in  
 CC electronic format directly from WIPO at  
 CC ftp.wipo.int/pub/published\_pct\_sequences

XX SQ Sequence 729 BP; 218 A; 116 C; 184 G; 211 T; 0 U; 0 Other;

Query Match 1.7%; Score 21; DB 7; Length 729;  
 Best Local Similarity 100.0%; Pred. No. 7.8;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 22 AAGATGAAGATGTTTGATA 42  
 Db 290 AAGATGAAGATGTTTGATA 310

RESULT 5  
 ABQ70805/c  
 ID ABQ70805 standard; DNA; 2791 BP.

AC ABQ70805;

XX 29-AUG-2003 (revised)

DT 29-AUG-2002 (first entry)

XX Listeria monocytogenes 4b contig DNA sequence #747.

XX Antibacterial; *Listeria*; food contamination; mutational analysis;  
 infection; ds.

XX Listeria monocytogenes ATCC 19115.

XX WO200228891-A2.

XX 11-APR-2002.

XX 04-OCT-2001; 2001WO-FR003061.

XX 04-OCT-2000; 2000FR-00012697.

XX (INSP ) INST PASTEUR.

PA (CNRS ) CNRS CENT NAT RECH SCI.

XX Kunst F, Glaser P;

XX WPI; 2002-332479/37.

XX New genomic sequences from *Listeria* species, useful for detection,  
 PT treatment and prevention of infection, also related polypeptides,  
 PT antibodies and modulators.

XX Claim 14; SEQ ID NO 3618; 180pp; French.

XX The present invention relates to nucleic acid sequences (ABQ67188-  
 CC ABQ71212) from *Listeria* sp. The sequences are useful as probes and  
 CC primers for identification and/or detection of *Listeria* (e.g. as  
 CC contaminants in foods, or mutational analysis) and for analysis of gene  
 CC expression. Proteins encoded by the nucleic acid sequences can be used to  
 CC screen for compounds that modulate gene expression, replication and  
 CC pathogenicity of *Listeria* (potential therapeutic agents), also for  
 CC treating infections by *Listeria*, and are useful as immunogens in anti-  
 CC *Listeria* vaccines. Note: The sequence data for this patent did not form  
 CC part of the printed specification, but was obtained in electronic format  
 CC directly from WIPO at ftp.wipo.int/pub/published\_pct\_sequences. (Updated  
 CC on 29-AUG-2003 to standardise OS field)

XX SQ Sequence 2791 BP; 942 A; 450 C; 622 G; 777 T; 0 U; 0 Other;

Query Match 1.7%; Score 21; DB 6; Length 2791;  
 Best Local Similarity 100.0%; Pred. No. 7.5;  
 Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 685 GCTACTTTTTCGTGTCATT 705

Db 693 GCTACTTTTTCGTGTCATT 673

RESULT 6

AAK82159  
 ID AAK82159 standard; DNA; 25392 BP.

XX AAK82159;

DT 07-NOV-2001 (first entry)

XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:36971.

XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;  
 KW cytostatic; gene therapy; vaccine; metastasis; ds.

XX Homo sapiens.

XX WO200157182-A2.

XX 09-AUG-2001.

XX 17-JAN-2001; 2001WO-US001354.

XX 31-JAN-2000; 2000US-0179065P.

XX 04-FEB-2000; 2000US-0180628P.

XX 24-FEB-2000; 2000US-0184664P.

XX 02-MAR-2000; 2000US-0186350P.

XX 16-MAR-2000; 2000US-0189874P.

XX 17-MAR-2000; 2000US-0190076P.

XX 18-APR-2000; 2000US-0198123P.

XX 19-MAY-2000; 2000US-0205515P.

XX 07-JUN-2000; 2000US-0209467P.

XX 28-JUN-2000; 2000US-0214886P.

XX 30-JUN-2000; 2000US-0215135P.

XX 07-JUL-2000; 2000US-0216880P.

XX 11-JUL-2000; 2000US-0217487P.

XX 14-JUL-2000; 2000US-0217496P.

XX 26-JUL-2000; 2000US-0218290P.

XX 26-JUL-2000; 2000US-0220963P.

XX 14-AUG-2000; 2000US-0220964P.

XX 14-AUG-2000; 2000US-0224518P.

XX 14-AUG-2000; 2000US-0224519P.

XX 14-AUG-2000; 2000US-0225213P.

XX 14-AUG-2000; 2000US-0225214P.

XX 14-AUG-2000; 2000US-0225266P.

XX 14-AUG-2000; 2000US-0225267P.

XX 14-AUG-2000; 2000US-0225268P.

XX 14-AUG-2000; 2000US-0225270P.

XX 14-AUG-2000; 2000US-0225447P.

PR	14-AUG-2000;	2000US-0225757P.
PR	14-AUG-2000;	2000US-0225758P.
PR	14-AUG-2000;	2000US-0225759P.
PR	18-AUG-2000;	2000US-0226279P.
PR	22-AUG-2000;	2000US-0226681P.
PR	22-AUG-2000;	2000US-0226868P.
PR	22-AUG-2000;	2000US-0227182P.
PR	23-AUG-2000;	2000US-0227009P.
PR	30-AUG-2000;	2000US-0228924P.
PR	01-SEP-2000;	2000US-0229343P.
PR	01-SEP-2000;	2000US-0229344P.
PR	01-SEP-2000;	2000US-0229345P.
PR	05-SEP-2000;	2000US-0229513P.
PR	06-SEP-2000;	2000US-0230437P.
PR	06-SEP-2000;	2000US-0230438P.
PR	08-SEP-2000;	2000US-0231242P.
PR	08-SEP-2000;	2000US-0231243P.
PR	08-SEP-2000;	2000US-0231244P.
PR	08-SEP-2000;	2000US-0231413P.
PR	08-SEP-2000;	2000US-0231414P.
PR	08-SEP-2000;	2000US-0232080P.
PR	08-SEP-2000;	2000US-0232081P.
PR	12-SEP-2000;	2000US-0233968P.
PR	14-SEP-2000;	2000US-0233977P.
PR	14-SEP-2000;	2000US-0233988P.
PR	14-SEP-2000;	2000US-0234001P.
PR	14-SEP-2000;	2000US-0234011P.
PR	14-SEP-2000;	2000US-0233063P.
PR	14-SEP-2000;	2000US-0233064P.
PR	14-SEP-2000;	2000US-0233065P.
PR	21-SEP-2000;	2000US-0234223P.
PR	21-SEP-2000;	2000US-0234274P.
PR	25-SEP-2000;	2000US-0234597P.
PR	25-SEP-2000;	2000US-0234988P.
PR	26-SEP-2000;	2000US-0235484P.
PR	27-SEP-2000;	2000US-0235834P.
PR	27-SEP-2000;	2000US-0235836P.
PR	29-SEP-2000;	2000US-0236327P.
PR	29-SEP-2000;	2000US-0236367P.
PR	29-SEP-2000;	2000US-0236368P.
PR	29-SEP-2000;	2000US-0236369P.
PR	29-SEP-2000;	2000US-0236370P.
PR	02-OCT-2000;	2000US-0236802P.
PR	02-OCT-2000;	2000US-0237037P.
PR	02-OCT-2000;	2000US-0237038P.
PR	02-OCT-2000;	2000US-0237039P.
PR	02-OCT-2000;	2000US-0237040P.
PR	13-OCT-2000;	2000US-0239935P.
PR	13-OCT-2000;	2000US-0239937P.
PR	20-OCT-2000;	2000US-0240960P.
PR	20-OCT-2000;	2000US-0241221P.
PR	20-OCT-2000;	2000US-0241785P.
PR	20-OCT-2000;	2000US-0241786P.
PR	20-OCT-2000;	2000US-0241787P.
PR	20-OCT-2000;	2000US-0241808P.
PR	20-OCT-2000;	2000US-0241809P.
PR	01-NOV-2000;	2000US-0241826P.
PR	08-NOV-2000;	2000US-0246474P.
PR	08-NOV-2000;	2000US-0246475P.
PR	08-NOV-2000;	2000US-0246476P.
PR	08-NOV-2000;	2000US-0246477P.
PR	08-NOV-2000;	2000US-0246478P.
PR	08-NOV-2000;	2000US-0246523P.
PR	08-NOV-2000;	2000US-0246524P.
PR	08-NOV-2000;	2000US-0246525P.
PR	08-NOV-2000;	2000US-0246526P.
PR	08-NOV-2000;	2000US-0246527P.
PR	08-NOV-2000;	2000US-0246528P.
PR	08-NOV-2000;	2000US-0246532P.
PR	08-NOV-2000;	2000US-0246533P.
PR	08-NOV-2000;	2000US-0246609P.
PR	08-NOV-2000;	2000US-0246610P.
PR	08-NOV-2000;	2000US-0246611P.
PR	08-NOV-2000;	2000US-0246613P.
PR	17-NOV-2000;	2000US-0249207P.
PR	17-NOV-2000;	2000US-0249208P.
PR	17-NOV-2000;	2000US-0249209P.
PR	17-NOV-2000;	2000US-0249210P.
PR	17-NOV-2000;	2000US-0249211P.
PR	17-NOV-2000;	2000US-0249212P.
PR	17-NOV-2000;	2000US-0249213P.
PR	17-NOV-2000;	2000US-0249214P.
PR	17-NOV-2000;	2000US-0249215P.
PR	17-NOV-2000;	2000US-0249216P.
PR	17-NOV-2000;	2000US-0249217P.
PR	17-NOV-2000;	2000US-0249218P.
PR	17-NOV-2000;	2000US-0249244P.
PR	17-NOV-2000;	2000US-0249245P.
PR	17-NOV-2000;	2000US-0249264P.
PR	17-NOV-2000;	2000US-0249265P.
PR	17-NOV-2000;	2000US-0249297P.
PR	17-NOV-2000;	2000US-0249299P.
PR	17-NOV-2000;	2000US-0249300P.
PR	01-DEC-2000;	2000US-0250160P.
PR	01-DEC-2000;	2000US-0250391P.
PR	05-DEC-2000;	2000US-0251030P.
PR	05-DEC-2000;	2000US-0251988P.
PR	05-DEC-2000;	2000US-0256719P.
PR	06-DEC-2000;	2000US-0251479P.
PR	06-DEC-2000;	2000US-0251856P.
PR	08-DEC-2000;	2000US-0251868P.
PR	08-DEC-2000;	2000US-0251869P.
PR	08-DEC-2000;	2000US-0251989P.
PR	08-DEC-2000;	2000US-0251990P.
PR	11-DEC-2000;	2000US-0254037P.
PR	05-JAN-2001;	2001US-0259678P.
XX		
PA	(HUMA-) HUMAN GENOME SCI INC.	
XX		
PI	Rosen CA, Barash SC, Ruben SM;	
XX	WPI; 2001-483426/52.	
DR		
XX		
PT	Nucleic acids encoding human immune/hematopoietic antigen polypeptides,	
PT	useful for preventing, diagnosing and/or treating cancers and metastasis.	
XX	Disclosure; SEQ ID NO 36971; 3071pp + Sequence Listing; English.	
PS		
XX	AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)	
CC	amino acid sequences given in AAK82170 to AAK91921. (I) have cytostatic	
CC	activity, and can be used in gene therapy and vaccine production. (I)	
CC	proteins and polynucleotides may be used in the prevention, diagnosis and	
CC	treatment of diseases associated with inappropriate (I) expression. For	
CC	example, they may be used to treat disorders associated with decreased	
CC	expression by rectifying mutations or deletions in a patient's genome	
CC	that affect the activity of (I) by expressing inactive proteins or to	
CC	supplement the patient's own production of (I). Additionally, (I)	
CC	polynucleotides may be used to produce the secreted (I), by inserting the	
CC	nucleic acids into a host cell and culturing the cell to express the	
CC	protein. (I) proteins and polynucleotides may be used to prevent,	
CC	diagnose and treat immune/haematopoietic-related diseases, especially	
CC	cancers and cancer metastases of haematopoietic-derived cells. AAK64703	
CC	to AAK87694 represent human immune/haematopoietic antigen genomic	
CC	sequences from the present invention. AAK54942 to AAK54950 and AAK82169	
CC	represent sequences used in the exemplification of the present invention	
XX		
SQ	Sequence 25392 BP; 7641 A; 4435 C; 4734 G; 8582 T; 0 U; 0 Other;	
	Query Match 1.7%; Score 21; DB 4; Length 25392;	
	Best Local Similarity 100.0%; Pred. No. 6.9;	
	Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;	
QY	1164 TTGTGTTGTTTGTATTGTT 1184	



```
Db 21789 TTGGTGTGTTGTAATGTT 21809
|||||
RESULT 7
ABA03041_20
Continuation (21 of 30) of ABA03041 from base 2000001 (Listeria monocytogenes EGD-e gene)
WP Sequence split into 30 fragments LOCUS ABA03041 Accession ABA03041
WP Fragment Name Begin End
WP ABA03041_00 1 110000
WP ABA03041_01 100001 210000
WP ABA03041_02 200001 310000
WP ABA03041_03 300001 410000
WP ABA03041_04 400001 510000
WP ABA03041_05 500001 610000
WP ABA03041_06 600001 710000
WP ABA03041_07 700001 810000
WP ABA03041_08 800001 910000
WP ABA03041_09 900001 1010000
WP ABA03041_10 1000001 1110000
WP ABA03041_11 1100001 1210000
WP ABA03041_12 1200001 1310000
WP ABA03041_13 1300001 1410000
WP ABA03041_14 1400001 1510000
WP ABA03041_15 1500001 1610000
WP ABA03041_16 1600001 1710000
WP ABA03041_17 1700001 1810000
WP ABA03041_18 1800001 1910000
WP ABA03041_19 1900001 2010000
WP ABA03041_20 2000001 2110000
WP ABA03041_21 2100001 2210000
WP ABA03041_22 2200001 2310000
WP ABA03041_23 2300001 2410000
WP ABA03041_24 2400001 2510000
WP ABA03041_25 2500001 2610000
WP ABA03041_26 2600001 2710000
WP ABA03041_27 2700001 2810000
WP ABA03041_28 2800001 2910000
WP ABA03041_29 2900001 2944528

Query Match 1.7%; Score 21; DB 6; Length 110000;
Best Local Similarity 100.0%; Pred. No. 6.6;
Matches 21; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 685 GCTACTTTTCTGTGTCAT 705
|||||
Db 11086 GCTACTTTTCTGTGTCAT 11106

RESULT 8
ACA36531/c
ID ACA36531 standard; DNA; 1017 BP.
AC ACA36531;
XX
XX 19-JUN-2003 (first entry)
XX Prokaryotic essential gene #18188.
XX
XX Antisense; ds; prokaryotic essential gene; cell proliferation;
XX drug design; gene.
XX Listeria monocytogenes.
XX
XX WO200277183-A2.
XX
XX 03-OCT-2002.
XX
XX 21-MAR-2002; 2002WO-US009107.
XX
XX 21-MAR-2001; 2001US-00815242.
XX
XX 06-SEP-2001; 2001US-00948993.
XX
XX 25-OCT-2001; 2001US-0342923P.
```

```
PR 08-FEB-2002; 2002US-00072851.
XX 06-MAR-2002; 2002US-0362699P.
XX (ELIT-) ELITRA PHARM INC.
XX
XX Wang L, Zamudio C, Malone C, Haselbeck R, Ohlsen KL, Zyskind JW;
XX Wall D, Trawick JD, Carr GJ, Yamamoto R, Forsyth RA, Xu HH;
XX
XX WPI; 2003-029926/02.
XX P-PSDB; ABU32661.
XX
XX New antisense nucleic acids, useful for identifying proteins or screening
XX for homologous nucleic acids required for cellular proliferation to
XX isolate candidate molecules for rational drug discovery programs.
XX
XX Claim 14; SEQ ID NO 24401; 17666pp; English.
XX
XX The invention relates to an isolated nucleic acid comprising any one of
XX the 6213 antisense sequences given in the specification where expression
XX of the nucleic acid inhibits proliferation of a cell. Also included are:
XX (1) a vector comprising a promoter operably linked to the nucleic acid
XX encoding a polypeptide whose expression is inhibited by the antisense
XX nucleic acid; (2) a host cell containing the vector; (3) an isolated
XX polypeptide or its fragment whose expression is inhibited by the
XX antisense nucleic acid; (4) an antibody capable of specifically binding
XX the polypeptide; (5) producing the polypeptide; (6) inhibiting cellular
XX proliferation or the activity of a gene in an operon required for
XX proliferation; (7) identifying a compound that influences the activity of
XX the gene product or that has an activity against a biological pathway
XX required for proliferation, or that inhibits cellular proliferation; (8)
XX identifying a gene required for cellular proliferation or the biological
XX pathway in which a proliferation-required gene or its gene product lies
XX or a gene on which the test compound that inhibits proliferation of an
XX organism acts; (9) manufacturing an antibiotic; (10) profiling a
XX compound's activity; (11) a culture comprising strains in which the gene
XX product is overexpressed or underexpressed; (12) determining the extent
XX to which each of the strains is present in a culture or collection of
XX strains; or (13) identifying the target of a compound that inhibits the
XX proliferation of an organism. The antisense nucleic acids are useful for
XX identifying proteins or screening for homologous nucleic acids required
XX for cellular proliferation to isolate candidate molecules for rational
XX drug discovery programs, or for screening homologous nucleic acids
XX required for proliferation in cells other than S. aureus, S. typhimurium,
XX K. pneumoniae or P. aeruginosa. The present sequence is one of the target
XX prokaryotic essential genes. Note: The sequence data for this patent did
XX not form part of the printed specification, but was obtained in
XX electronic format directly from WIPO at
XX ftp.wipo.int/pub/published_pct_sequences
XX
XX SQ Sequence 1017 BP; 326 A; 170 C; 237 G; 284 T; 0 U; 0 Other;
```

```
Query Match 1.6%; Score 20; DB 7; Length 1017;
Best Local Similarity 100.0%; Pred. No. 24;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 685 GCTACTTTTCTGTGTCAT 704
|||||
Db 20 GCTACTTTTCTGTGTCAT 1

RESULT 9
AAD25856
ID AAD25856 standard; DNA; 16488 BP.
XX
XX AAD25856;
XX
XX 26-MAR-2002 (first entry)
XX Human diphtheria toxin receptor (DTR) gene.
XX
XX Human; polymorphic site; PS; diphtheria toxin receptor; DTR; haplotype;
XX heparin-binding epidermal growth factor-like growth factor; therapy;
XX chromosome 5q23; transgenic animal; drug screening; tumour growth;
```

useful for screening or expressing proteins for treating diseases related to DTR activity.

Claim 21; Fig 1; 66pp; English.

The present invention relates to an isolated polynucleotide, comprising polymorphisms in the human diphtheria toxin receptor, heparin-binding epidermal growth factor-like growth factor (DTR) gene. DTR gene is located on chromosome 5q23. The polynucleotide comprising polymorphisms in the DTR gene is useful in studying the expression and function of DTR, and in expressing DTR protein for use in screening candidate drugs to treat diseases related to DTR activity. The methods and haplotypes are useful in improving the efficiency and output of several steps in the drug discovery and development process, including target validation, identifying lead compounds, and early phase clinical trials. The kit and method are useful for determining if an individual has one of the haplotypes or haplotype pairs. The transgenic animals are useful for studying expression of the DTR isogenes in vivo, for in vivo screening and testing of drugs targeted against DTR protein, and for testing the efficacy of therapeutic agents and compounds for tumour growth, smooth muscle hyperplasia or atherosclerosis in a biological system. The present sequence is DTR gene

Sequence 16488 BP; 3925 A; 3939 C; 4220 G; 4404 T; 0 U; 0 Other;

Query Match 1.6%; Score 20; DB 6; Length 16488;  
 Best Local Similarity 100.0%; Pred. No. 22;  
 Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 869 CTTCTGGTGTCTGATAAGT 888  
 |||||  
 Db 16270 CTTCTGGTGTCTGATAAGT 16289

RESULT 10  
 AAD25895  
 ID AAD25895 standard; DNA; 16488 BP.  
 XX AAD25895;  
 AC AAD25895;  
 XX 26-MAR-2002 (first entry)  
 DT Human diphtheria toxin receptor (DTR) gene allelic variant.  
 DE Human; polymorphic site; PS; diphtheria toxin receptor; DTR; haplotype;  
 KW heparin-binding epidermal growth factor-like growth factor; therapy;  
 KW chromosome 5q23; transgenic animal; drug screening; tumour growth;  
 KW smooth muscle hyperplasia; atherosclerosis; ds.  
 XX Homo sapiens.  
 OS  
 FH Key Location/Qualifiers  
 FT misc\_feature /tag= a  
 FT /note= "This degenerate base represents polymorphic site,  
 FT PS1"  
 FT misc\_feature /tag= b  
 FT /note= "This degenerate base represents polymorphic site,  
 FT PS2"  
 FT misc\_feature /tag= c  
 FT /note= "This degenerate base represents polymorphic site,  
 FT PS3"  
 FT misc\_feature /tag= d  
 FT /note= "This degenerate base represents polymorphic site,  
 FT PS4"  
 FT misc\_feature /tag= e  
 FT /note= "This degenerate base represents polymorphic site,  
 FT PS5"  
 FT misc\_feature 11323

smooth muscle hyperplasia; atherosclerosis; ds.  
 Homo sapiens.

Key Location/Qualifiers  
 replace(737, C)  
 /tag= a  
 /note= "Polymorphic site, PS1"  
 replace(788, T)  
 /tag= b  
 /note= "Polymorphic site, PS2"  
 836..12488  
 /tag= c  
 /product= "Human DTR protein"  
 836..881  
 /tag= d  
 /number= 1  
 882..1079  
 /tag= e  
 replace(938, A)  
 /tag= f  
 /note= "Polymorphic site, PS3"  
 1080..1253  
 /tag= g  
 /number= 2  
 1254..14351  
 /tag= h  
 4352..4529  
 /tag= i  
 /number= 3  
 replace(4404, T)  
 /tag= j  
 /note= "Polymorphic site, PS4"  
 replace(4527, T)  
 /tag= k  
 /note= "Polymorphic site, PS5"  
 4530..11136  
 /tag= l  
 11137..11292  
 /tag= m  
 /number= 4  
 11293..12415  
 /tag= n  
 replace(11323, A)  
 /tag= o  
 /note= "Polymorphic site, PS6"  
 replace(11415, A)  
 /tag= p  
 /note= "Polymorphic site, PS7"  
 replace(12392, G)  
 /tag= q  
 /note= "Polymorphic site, PS8"  
 12416..12488  
 /tag= x  
 /number= 5

WO200179233-A2.  
 25-OCT-2001.  
 16-APR-2001; 2001WO-US012302.  
 14-APR-2000; 2000US-0197375P.  
 (GENA-) GENAISSANCE PHARM INC.  
 Choi JY, Kliem SE, Koshy B, Parks KE, Stephens JC;  
 WPI; 2002-082745/11.  
 P-PSDB; AAE15908.  
 New nucleotide polymorphisms in the human diphtheria toxin receptor,  
 heparin-binding epidermal growth factor-like growth factor (DTR) gene,

```
PT /*tag= f
FT /note= "This degenerate base represents polymorphic site,
PS6"
FT misc_feature
FT 11415
FT /*tag= g
FT /note= "This degenerate base represents polymorphic site,
PS7"
FT misc_feature
FT 12392
FT /*tag= h
FT /note= "This degenerate base represents polymorphic site,
PS8"
FT WO200179233-A2.
XX 25-OCT-2001.
XX 16-APR-2001; 2001WO-US012302.
XX 14-APR-2000; 2000US-0197375P.
XX (GENA-) GENAISSANCE PHARM INC.
XX Choi JY, Kliem SE, Koshy B, Parks KE, Stephens JC;
XX WPI; 2002-082745/11.
XX
XX New nucleotide polymorphisms in the human diphtheria toxin receptor,
FT heparin-binding epidermal growth factor-like growth factor (DTR) gene,
FT useful for screening or expressing proteins for treating diseases related
FT to DTR activity.
XX
XX Disclosure; Page 60-66; 66pp; English.
XX
XX The present invention relates to an isolated polynucleotide, comprising
CC polymorphisms in the human diphtheria toxin receptor, heparin-binding
CC epidermal growth factor-like growth factor (DTR) gene. DTR gene is
CC located on chromosome 5q22. The polynucleotide comprising polymorphisms
CC in the DTR gene is useful in studying the expression and function of DTR,
CC and in expressing DTR protein for use in screening candidate drugs to
CC treat diseases related to DTR activity. The methods and haplotypes are
CC useful in improving the efficiency and output of several steps in the
CC drug discovery and development process, including target validation,
CC identifying lead compounds, and early phase clinical trials. The kit and
CC method are useful for determining if an individual has one of the
CC haplotypes or haplotype pairs. The transgenic animals are useful for
CC studying expression of the DTR isogenes in vivo, for in vivo screening
CC and testing of drugs targeted against DTR protein, and for testing the
CC efficacy of therapeutic agents and compounds for tumour growth, smooth
CC muscle hyperplasia or atherosclerosis in a biological system. The present
CC sequence is human DTR gene allelic variant
XX
SQ Sequence 16488 BP; 3924 A; 3935 C; 4217 G; 4404 T; 0 U; 8 Other;
Query Match 1.68; Score 20; DB 6; Length 16488;
Best Local Similarity 100.0%; Pred. No. 22;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 869 CTTCTGCTGTTCTGATAAGT 888
DB 16270 CTTCTGCTGTTCTGATAAGT 16289
RESULT 11
AAK68992/c
ID AAK68992 standard; DNA; 42738 BP.
XX AAK68992;
XX
XX 06-NOV-2001 (first entry)
XX Human immune/haematopoietic antigen genomic sequence SEQ ID NO:23804.
XX Human; immune; haematopoietic; immune/haematopoietic antigen; cancer;
KW
```

```
KW cytostatic; gene therapy; vaccine; metastasis; ds.
XX Homo sapiens.
OS
XX WO200157182-A2.
PN
XX 09-AUG-2001.
PD
XX
XX 17-JAN-2001; 2001WO-US001354.
XX
XX 31-JAN-2000; 2000US-0179065P.
XX 04-FEB-2000; 2000US-0180628P.
XX 24-FEB-2000; 2000US-0184664P.
XX 02-MAR-2000; 2000US-0186350P.
XX 16-MAR-2000; 2000US-0189874P.
XX 17-MAR-2000; 2000US-0190076P.
XX 18-APR-2000; 2000US-0198123P.
XX 19-MAY-2000; 2000US-0205515P.
XX 07-JUN-2000; 2000US-0209467P.
XX 28-JUN-2000; 2000US-0214886P.
XX 30-JUN-2000; 2000US-0215135P.
XX 07-JUL-2000; 2000US-0216647P.
XX 07-JUL-2000; 2000US-0216880P.
XX 11-JUL-2000; 2000US-0217487P.
XX 11-JUL-2000; 2000US-0217496P.
XX 14-JUL-2000; 2000US-0218290P.
XX 26-JUL-2000; 2000US-0220963P.
XX 26-JUL-2000; 2000US-0220964P.
XX 14-AUG-2000; 2000US-0224518P.
XX 14-AUG-2000; 2000US-0224519P.
XX 14-AUG-2000; 2000US-0225213P.
XX 14-AUG-2000; 2000US-0225214P.
XX 14-AUG-2000; 2000US-0225267P.
XX 14-AUG-2000; 2000US-0225266P.
XX 14-AUG-2000; 2000US-0225268P.
XX 14-AUG-2000; 2000US-0225270P.
XX 14-AUG-2000; 2000US-0225447P.
XX 14-AUG-2000; 2000US-0225757P.
XX 14-AUG-2000; 2000US-0225758P.
XX 14-AUG-2000; 2000US-0225759P.
XX 18-AUG-2000; 2000US-0226279P.
XX 22-AUG-2000; 2000US-0226681P.
XX 22-AUG-2000; 2000US-0226868P.
XX 22-AUG-2000; 2000US-0227182P.
XX 23-AUG-2000; 2000US-0227009P.
XX 30-AUG-2000; 2000US-0228924P.
XX 01-SEP-2000; 2000US-0229287P.
XX 01-SEP-2000; 2000US-0229343P.
XX 01-SEP-2000; 2000US-0229344P.
XX 01-SEP-2000; 2000US-0229345P.
XX 05-SEP-2000; 2000US-0229509P.
XX 05-SEP-2000; 2000US-0229513P.
XX 06-SEP-2000; 2000US-0230437P.
XX 06-SEP-2000; 2000US-0230438P.
XX 08-SEP-2000; 2000US-0231242P.
XX 08-SEP-2000; 2000US-0231243P.
XX 08-SEP-2000; 2000US-0231244P.
XX 08-SEP-2000; 2000US-0231413P.
XX 08-SEP-2000; 2000US-0231414P.
XX 08-SEP-2000; 2000US-0232080P.
XX 12-SEP-2000; 2000US-0232081P.
XX 14-SEP-2000; 2000US-0231968P.
XX 14-SEP-2000; 2000US-0232197P.
XX 14-SEP-2000; 2000US-0232398P.
XX 14-SEP-2000; 2000US-0232399P.
XX 14-SEP-2000; 2000US-0232400P.
XX 14-SEP-2000; 2000US-0232401P.
XX 14-SEP-2000; 2000US-0233063P.
XX 14-SEP-2000; 2000US-0233064P.
XX 14-SEP-2000; 2000US-0233065P.
XX 21-SEP-2000; 2000US-0234223P.
XX 21-SEP-2000; 2000US-0234274P.
XX 25-SEP-2000; 2000US-0234997P.
```

Mon Sep 27 08:51:12 2004

```

PR 25-SEP-2000; 2000US-0234998P.
PR 26-SEP-2000; 2000US-0235484P.
PR 27-SEP-2000; 2000US-0235834P.
PR 27-SEP-2000; 2000US-0235836P.
PR 29-SEP-2000; 2000US-0236327P.
PR 29-SEP-2000; 2000US-0236367P.
PR 29-SEP-2000; 2000US-0236368P.
PR 29-SEP-2000; 2000US-0236369P.
PR 29-SEP-2000; 2000US-0236370P.
PR 02-OCT-2000; 2000US-0236802P.
PR 02-OCT-2000; 2000US-0237037P.
PR 02-OCT-2000; 2000US-0237038P.
PR 02-OCT-2000; 2000US-0237039P.
PR 13-OCT-2000; 2000US-0237040P.
PR 13-OCT-2000; 2000US-0239935P.
PR 13-OCT-2000; 2000US-0239937P.
PR 20-OCT-2000; 2000US-0240960P.
PR 20-OCT-2000; 2000US-0241221P.
PR 20-OCT-2000; 2000US-0241785P.
PR 20-OCT-2000; 2000US-0241786P.
PR 20-OCT-2000; 2000US-0241787P.
PR 20-OCT-2000; 2000US-0241808P.
PR 20-OCT-2000; 2000US-0241809P.
PR 20-OCT-2000; 2000US-0241826P.
PR 01-NOV-2000; 2000US-0244617P.
PR 01-NOV-2000; 2000US-0246474P.
PR 08-NOV-2000; 2000US-0246475P.
PR 08-NOV-2000; 2000US-0246476P.
PR 08-NOV-2000; 2000US-0246477P.
PR 08-NOV-2000; 2000US-0246478P.
PR 08-NOV-2000; 2000US-0246523P.
PR 08-NOV-2000; 2000US-0246524P.
PR 08-NOV-2000; 2000US-0246525P.
PR 08-NOV-2000; 2000US-0246526P.
PR 08-NOV-2000; 2000US-0246527P.
PR 08-NOV-2000; 2000US-0246528P.
PR 08-NOV-2000; 2000US-0246532P.
PR 08-NOV-2000; 2000US-0246609P.
PR 08-NOV-2000; 2000US-0246610P.
PR 08-NOV-2000; 2000US-0246611P.
PR 08-NOV-2000; 2000US-0246613P.
PR 17-NOV-2000; 2000US-0249207P.
PR 17-NOV-2000; 2000US-0249208P.
PR 17-NOV-2000; 2000US-0249209P.
PR 17-NOV-2000; 2000US-0249210P.
PR 17-NOV-2000; 2000US-0249211P.
PR 17-NOV-2000; 2000US-0249212P.
PR 17-NOV-2000; 2000US-0249213P.
PR 17-NOV-2000; 2000US-0249214P.
PR 17-NOV-2000; 2000US-0249215P.
PR 17-NOV-2000; 2000US-0249216P.
PR 17-NOV-2000; 2000US-0249217P.
PR 17-NOV-2000; 2000US-0249218P.
PR 17-NOV-2000; 2000US-0249244P.
PR 17-NOV-2000; 2000US-0249245P.
PR 17-NOV-2000; 2000US-0249246P.
PR 17-NOV-2000; 2000US-0249255P.
PR 17-NOV-2000; 2000US-0249257P.
PR 17-NOV-2000; 2000US-0249259P.
PR 17-NOV-2000; 2000US-0249300P.
PR 01-DEC-2000; 2000US-0250160P.
PR 01-DEC-2000; 2000US-0250391P.
PR 05-DEC-2000; 2000US-0251030P.
PR 05-DEC-2000; 2000US-0251988P.
PR 06-DEC-2000; 2000US-0256719P.
PR 06-DEC-2000; 2000US-0251479P.
PR 08-DEC-2000; 2000US-0251856P.
PR 08-DEC-2000; 2000US-0251868P.
PR 08-DEC-2000; 2000US-0251869P.
PR 08-DEC-2000; 2000US-0251989P.
PR 08-DEC-2000; 2000US-0251990P.
PR 11-DEC-2000; 2000US-0254097P.
PR 05-JAN-2001; 2001US-0259678P.

XX (HUMA-) HUMAN GENOME SCI INC.
XX Rosen CA, Barash SC, Ruben SM;
XX WPI; 2001-483426/52.
XX Nucleic acids encoding human immune/hematopoietic antigen polypeptides,
XX useful for preventing, diagnosing and/or treating cancers and metastasis.
XX Disclosure; SEQ ID NO 23804; 307lpp + Sequence Listing; English.
XX AAK54951 to AAK64702 encode the human immune/haematopoietic antigen (I)
XX amino acid sequences given in AAM82170 to AAM91921. (I) have cytostatic
XX activity, and can be used in gene therapy and vaccine production. (I)
XX proteins and polynucleotides may be used in the prevention, diagnosis and
XX treatment of diseases associated with inappropriate (I) expression. For
XX example, they may be used to treat disorders associated with decreased
XX expression by rectifying mutations or deletions in a patient's genome
XX that affect the activity of (I) by expressing inactive proteins or to
XX supplement the patients own production of (I). Additionally, (I)
XX polynucleotides may be used to produce the secreted (I), by inserting the
XX nucleic acids into a host cell and culturing the cell to express the
XX protein. (I) proteins and polynucleotides may be used to prevent,
XX diagnose and treat immune/haematopoietic-related diseases, especially
XX cancers and cancer metastases of haematopoietic-derived cells. AAK64703
XX to AAK87694 represent human immune/haematopoietic antigen genomic
XX sequences from the present invention. AAK54942 to AAK54950 and AAM82169
XX represent sequences used in the exemplification of the present invention
XX Sequence 42738 BP; 15539 A; 8490 C; 8408 G; 10301 T; 0 U; 0 Other;
Query Match 1.6%; Score 20; DB 4; Length 42738;
Best Local Similarity 100.0%; Pred. No. 21;
Matches 20; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1194 AAGTATGAATAAAATTGAA 1213
| | | | | | | | | | | | | | | |
DB 36375 AAGTATGAATAAAATTGAA 36356

RESULT 12
ADB95081
ID ADB95081 standard; DNA; 435 BP.
XX ADB95081;
XX AC ADB95081;
XX DT 04-DEC-2003 (first entry)
XX DE A. thaliana gene 60944 #SEQ ID 79.
XX KW Plant; herbicide; weed; crop field; growth; development; gene; ds.
XX OS Arabidopsis thaliana.
XX PN WO2003008440-A2.
XX PD 30-JAN-2003.
XX PF 16-JUL-2002; 2002WO-EP007929.
XX PR 16-JUL-2001; 2001US-0305806P.
XX PR 20-FEB-2002; 2002US-0358416P.
XX (SYGN ) SYNGENTA PARTICIPATIONS AG.
XX Levin JZ, Patton DA, McElver JA, Budziszewski GJ, Zhou Q, Aux GW;
XX Tossberg J, Wegrich Glover L, Ashby CS, Thomas CR, Madhaven E;
XX Lewis S, Dunn J, Cates E, Law MD;
XX WPI; 2003-229557/22.
XX DR P-PSDB; ADB95082.
XX
```

PT Identifying an herbicidal compound, useful for controlling undesirable  
PT vegetation, comprises combining a polypeptide with a compound to be  
PT tested for the ability to bind to the polypeptide or inhibit the activity  
XX of the polypeptide.

XX Example 9; SEQ ID NO 79; 273pp; English.

XX The invention relates to a method for identifying a herbicidal compound.  
CC The method of the invention comprises combining a polypeptide having at  
CC least 90% identical to any one of 48 69-1008 residue amino acid sequences  
CC (designated as P1-P48), given in the specification, with a compound to be  
CC tested for the ability to bind to the polypeptide or inhibit the activity  
CC of the polypeptide, under conditions conducive to binding or inhibiting,  
CC respectively. Also disclosed is a method for killing or inhibiting the  
CC growth or viability of a plant by applying to the plant the herbicidal  
CC compound identified by the novel method, a chimeric construct comprising  
CC a promoter operatively linked to the nucleic acid molecule, a recombinant  
CC vector comprising the chimeric construct and a host cell comprising the  
CC nucleic acid molecule. The method and polypeptides are useful in  
CC screening assays to identify compounds that interact or inhibit the  
CC polypeptides, thus as potential herbicides to control undesirable  
CC vegetation such as weeds in crop fields. Nucleic acid molecules (odd  
CC numbers between ADB95003 and ADB95097) isolated from Arabidopsis thaliana  
CC comprising nucleotide sequences that encode proteins (even numbers  
CC between ADB95004-ADB95098) are essential for plant growth and  
CC development.

XX SQ Sequence 435 BP; 130 A; 72 C; 122 G; 111 T; 0 U; 0 Other;

Query Match 1.6%; Score 19; DB 9; Length 435;

Best Local Similarity 100.0%; Pred. No. 78;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 532 AAGAAAGAGTATGTTGTTG 550

DB 229 AAGAAAGAGTATGTTGTTG 247

RESULT 13

ABQ38560

ID ABQ38560 standard; DNA; 524 BP.

XX AC ABQ38560;

XX DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 25151.

XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
XX KW drug; side effect; cancer; central nervous system; cardiovascular;  
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
XX KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX PT Determining the degree of cytosine methylation in genomic DNA, useful for  
XX PT diagnosis and prognosis, comprises selective hybridization of amplicons  
XX PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

XX This invention describes a novel method for determining the degree of  
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a  
CC genomic sample of DNA. The sample is treated chemically to convert  
CC cytosine (C) but not methylated C, to uracil, then part of the genomic  
CC DNA that contains the target C is amplified to form a labeled amplicon.  
CC The amplicon is hybridised to two classes, each with at least one member,  
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the  
CC degree of hybridisation to both classes is determined from the label on  
CC the amplicon. From the ratio of labels hybridised to the two classes of  
CC oligomers, the degree of methylation is calculated. The method is used:  
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs  
CC and of a wide range of diseases, e.g. cancer, disorders of the central  
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,  
CC particularly by detecting mutations or single nucleotide polymorphisms  
CC (SNPs); and (ii) for differentiation of cell or tissue types and for  
CC investigating cell differentiation. The method allows the methylation  
CC status of many C residues to be determined simultaneously. ABQ13410-  
CC ABQ54121 represent genomic DNA sequences used to illustrate the method  
CC for determining the degree of cytosine methylation described in the  
CC disclosure of the invention

XX SQ Sequence 524 BP; 26 A; 60 C; 174 G; 264 T; 0 U; 0 Other;

Query Match 1.6%; Score 19; DB 6; Length 524;

Best Local Similarity 100.0%; Pred. No. 78;

Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1165 TGTGTGTTTTGTATTGT 1183

DB 1 TGTGTGTTTTGTATTGT 19

RESULT 14

ABQ38561/c

ID ABQ38561 standard; DNA; 524 BP.

XX AC ABQ38561;

XX DT 12-JUL-2002 (first entry)

XX DE Oligonucleotide for detecting cytosine methylation SEQ ID NO 25152.

XX KW Human; cytosine methylation; 5'-CpG-3'; uracil; cytosine; diagnosis;  
XX KW drug; side effect; cancer; central nervous system; cardiovascular;  
XX KW gastrointestinal; respiratory system; single nucleotide polymorphism;  
XX KW SNP; cell differentiation; ds.

XX OS Homo sapiens.

XX PN WO200218632-A2.

XX PD 07-MAR-2002.

XX PF 01-SEP-2001; 2001WO-EP010074.

XX PR 01-SEP-2000; 2000DE-01043826.

XX PR 05-SEP-2000; 2000DE-01044543.

XX PA (EPIG-) EPIGENOMICS AG.

XX PI Olek A, Piepenbrock C, Berlin K, Guetig D;

XX DR WPI; 2002-371829/40.

XX PT Determining the degree of cytosine methylation in genomic DNA, useful for  
XX PT diagnosis and prognosis, comprises selective hybridization of amplicons  
XX PT from chemically treated DNA.

XX Claim 12; 56pp + Sequence Listing; 56pp; German.

Mon Sep 27 08:51:12 2004

us-09-913-767-1.oligo.rng

```
CC This invention describes a novel method for determining the degree of
CC methylation of a particular cytosine in a motif 5'-CpG-3', present in a
CC genomic sample of DNA. The sample is treated chemically to convert
CC cytosine (C) but not methylated C, to uracil, then part of the genomic
CC DNA that contains the target C is amplified to form a labeled amplicon.
CC The amplicon is hybridised to two classes, each with at least one member,
CC of oligonucleotides and/or peptide-nucleic acid (PNA) oligomers and the
CC degree of hybridisation to both classes is determined from the label on
CC the amplicon. From the ratio of labels hybridised to the two classes of
CC oligomers, the degree of methylation is calculated. The method is used:
CC (i) for diagnosis and/or prognosis of side effects of therapeutic drugs
CC and of a wide range of diseases, e.g. cancer, disorders of the central
CC nervous, cardiovascular, gastrointestinal and respiratory systems etc.,
CC particularly by detecting mutations or single nucleotide polymorphisms
CC (SNP's); and (ii) for differentiation of cell or tissue types and for
CC investigating cell differentiation. The method allows the methylation
CC status of many C residues to be determined simultaneously. ABQ13410-
CC ABQ54121 represent genomic DNA sequences used to illustrate the method
CC for determining the degree of cytosine methylation described in the
CC disclosure of the invention
CC
XX
SQ Sequence 524 BP; 264 A; 174 C; 60 G; 26 T; 0 U; 0 Other;

Query Match 1.6%; Score 19; DB 6; Length 524;
Best Local Similarity 100.0%; Pred. No. 78;
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1165 TGTGTGTTTTGTATTGT 1183
Db |||||||
524 TGTGTGTTTTGTATTGT 506

RESULT 15
AAC33529
ID AAC33529 standard; DNA; 644 BP.
XX
AC AAC33529;
XX
XX 17-OCT-2000 (first entry)
DT
DE
XX
XX Arabidopsis thaliana DNA fragment SEQ ID NO: 3385.
XX
XX Hybridisation assay; genetic mapping; gene expression control;
XX protein identification; signal transduction pathway; metabolic pathway;
XX promoter; termination sequence; ss.
XX
XX Arabidopsis thaliana.
XX
XX EP1033405-A2.
XX
XX 06-SEP-2000.
XX
XX 25-FEB-2000; 2000EP-00301439.
XX
XX 25-FEB-1999; 99US-0121825P.
XX 05-MAR-1999; 99US-0123180P.
XX 09-MAR-1999; 99US-0123548P.
XX 23-MAR-1999; 99US-0125788P.
XX 25-MAR-1999; 99US-0126264P.
XX 29-MAR-1999; 99US-0126785P.
XX 01-APR-1999; 99US-0127462P.
XX 06-APR-1999; 99US-0128234P.
XX 08-APR-1999; 99US-0128714P.
XX 16-APR-1999; 99US-0129845P.
XX 19-APR-1999; 99US-0130077P.
XX 21-APR-1999; 99US-0130449P.
XX 23-APR-1999; 99US-0130510P.
XX 28-APR-1999; 99US-0130891P.
XX 30-APR-1999; 99US-0131449P.
XX 30-APR-1999; 99US-0132048P.
XX 04-MAY-1999; 99US-0132407P.
XX 05-MAY-1999; 99US-0132484P.
XX 05-MAY-1999; 99US-0132485P.
XX 06-MAY-1999; 99US-0132486P.
XX 06-MAY-1999; 99US-0132487P.
XX 07-MAY-1999; 99US-0132863P.
XX 11-MAY-1999; 99US-0134256P.
XX 14-MAY-1999; 99US-0134218P.
XX 14-MAY-1999; 99US-0134219P.
XX 14-MAY-1999; 99US-0134221P.
XX 14-MAY-1999; 99US-0134370P.
XX 18-MAY-1999; 99US-0134768P.
XX 19-MAY-1999; 99US-0134941P.
XX 20-MAY-1999; 99US-0135124P.
XX 21-MAY-1999; 99US-0135353P.
XX 24-MAY-1999; 99US-0135629P.
XX 25-MAY-1999; 99US-0136021P.
XX 27-MAY-1999; 99US-0136392P.
XX 28-MAY-1999; 99US-0136782P.
XX 01-JUN-1999; 99US-0137222P.
XX 03-JUN-1999; 99US-0137528P.
XX 04-JUN-1999; 99US-0137528P.
XX 07-JUN-1999; 99US-0137724P.
XX 08-JUN-1999; 99US-0138094P.
XX 10-JUN-1999; 99US-0138540P.
XX 10-JUN-1999; 99US-0138847P.
XX 14-JUN-1999; 99US-0139119P.
XX 16-JUN-1999; 99US-0139452P.
XX 16-JUN-1999; 99US-0139453P.
XX 17-JUN-1999; 99US-0139492P.
XX 18-JUN-1999; 99US-0139454P.
XX 18-JUN-1999; 99US-0139455P.
XX 18-JUN-1999; 99US-0139456P.
XX 18-JUN-1999; 99US-0139457P.
XX 18-JUN-1999; 99US-0139458P.
XX 18-JUN-1999; 99US-0139459P.
XX 18-JUN-1999; 99US-0139460P.
XX 18-JUN-1999; 99US-0139461P.
XX 18-JUN-1999; 99US-0139462P.
XX 18-JUN-1999; 99US-0139463P.
XX 18-JUN-1999; 99US-0139750P.
XX 18-JUN-1999; 99US-0139763P.
XX 21-JUN-1999; 99US-0139817P.
XX 22-JUN-1999; 99US-0139899P.
XX 23-JUN-1999; 99US-0140353P.
XX 23-JUN-1999; 99US-0140354P.
XX 24-JUN-1999; 99US-0140695P.
XX 28-JUN-1999; 99US-0140823P.
XX 29-JUN-1999; 99US-0140991P.
XX 30-JUN-1999; 99US-0141287P.
XX 01-JUL-1999; 99US-0141842P.
XX 01-JUL-1999; 99US-0142154P.
XX 02-JUL-1999; 99US-0142055P.
XX 06-JUL-1999; 99US-0142390P.
XX 08-JUL-1999; 99US-0142803P.
XX 09-JUL-1999; 99US-0142920P.
XX 12-JUL-1999; 99US-0142977P.
XX 13-JUL-1999; 99US-0143542P.
XX 14-JUL-1999; 99US-0143624P.
XX 15-JUL-1999; 99US-0144005P.
XX 16-JUL-1999; 99US-0144085P.
XX 16-JUL-1999; 99US-0144086P.
XX 19-JUL-1999; 99US-0144325P.
XX 19-JUL-1999; 99US-0144331P.
XX 19-JUL-1999; 99US-0144332P.
XX 19-JUL-1999; 99US-0144333P.
XX 19-JUL-1999; 99US-0144334P.
XX 19-JUL-1999; 99US-0144335P.
XX 20-JUL-1999; 99US-0144352P.
XX 20-JUL-1999; 99US-0144632P.
XX 20-JUL-1999; 99US-0144884P.
XX 21-JUL-1999; 99US-0144814P.
XX 21-JUL-1999; 99US-0145086P.
XX 21-JUL-1999; 99US-0145088P.
XX 22-JUL-1999; 99US-0145085P.
XX 22-JUL-1999; 99US-0145087P.
```

PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145313P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 30-AUG-1999; 99US-0151080P.  
PR 31-AUG-1999; 99US-0151303P.  
PR 01-SEP-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.

PR 21-OCT-1999; 99US-0160814P.  
PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 23-OCT-1999; 99US-0161404P.  
PR 23-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161320P.  
PR 28-OCT-1999; 99US-0161992P.  
PR 28-OCT-1999; 99US-0161993P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 1.6%; Score 19; DB 3; Length 644;  
Best Local Similarity 100.0%; Pred. No. 77;  
Matches 19; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

Qy 532 AAGAAAGAGTATGTGGTTG 550  
|||||  
Db 311 AAGAAAGAGTATGTGGTTG 329  
|||||

Search completed: September 25, 2004, 11:49:19  
Job time : 612 secs

BEST AVAILABLE COPY

**THIS PAGE BLANK (USPTO)**



GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:33:00 ; Search time 89 Seconds  
(without alignments)  
1130.190 Million cell updates/sec

Title: US-09-913-767-8  
Perfect score: 1798  
Sequence: 1 MNGLIIINCIIITGCGG.....DRQPPELPILPVSQYVA 356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1586107 seqs, 282547505 residues

Total number of hits satisfying chosen parameters: 1586107

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : A\_Geneseq\_29Jan04:\*

- 1: Geneseqp1980s:\*
- 2: Geneseqp1990s:\*
- 3: Geneseqp2000s:\*
- 4: Geneseqp2001s:\*
- 5: Geneseqp2002s:\*
- 6: Geneseqp2003as:\*
- 7: Geneseqp2003bs:\*
- 8: Geneseqp2004s:\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1798	100.0	356	3	AAB10684
2	1792	99.7	356	5	ABB91157
3	1134.5	63.1	351	5	ABB91156
4	892	49.6	178	3	AAG34467
5	727	40.4	144	3	AAG15103
6	727	40.4	164	3	AAG15102
7	553	30.8	367	3	AAG22805
8	553	30.8	367	3	AAG40302
9	553	30.8	382	3	AAG22804
10	553	30.8	392	3	AAG40301
11	550	30.6	367	3	AAG44536
12	525	29.2	107	3	AAG15104
13	513.5	28.6	386	3	AAG52938
14	513.5	28.6	387	3	AAG52937
15	513.5	28.6	394	3	AAG52936
16	512.5	28.5	386	3	AAG16270
17	512.5	28.5	387	3	AAG16269
18	512.5	28.5	394	3	AAG16268
19	497	27.6	303	3	AAG44537
20	473.5	26.3	335	3	AAG17233
21	473.5	26.3	355	3	AAG17232
22	473.5	26.3	361	3	AAG17231
23	449.5	25.0	250	3	AAG44538
24	445	24.7	358	5	ABB93669
25	161	9.0	134	3	AAG16232

## ALIGNMENTS

## RESULT 1

AAB10684	
ID	AAB10684 standard; protein; 356 AA.
XX	
AC	AAB10684;
XX	
DT	19-JAN-2001 (first entry)
XX	
DE	A. thaliana PUP1 protein.
XX	
KW	PUP1; transgenic plant; nucleobase transporter; apical dominance; flowering behaviour; senescence; pesticide distribution.
KW	
XX	
OS	Arabidopsis thaliana.
XX	
PN	DE19907209-A1.
XX	
PD	24-AUG-2000.
XX	
PF	19-FEB-1999; 99DE-01007209.
XX	
PR	19-FEB-1999; 99DE-01007209.
XX	
PA	(FROM/) FROMMER W.
XX	
PI	Gillissen B, Buerkle L, Andre B, Frommer WB;
XX	
DR	WPI; 2000-566202/53.
DR	N-PSDB; AAA97919.
XX	
PT	Nucleic acid, useful for producing transgenic plants with altered nucleobase transport, encodes a nucleobase transporter protein of Arabidopsis thaliana.
XX	
PS	Claim 1b; Page 16-17; 24pp; German.
XX	
CC	This invention describes a novel nucleic acid encoding a plant nucleobase transporter (I). (I) is produced by complementation of a nucleobase transport (NBT)-defective host cell with a plant gene bank by selection of NBT-positive cells. (I) is used to isolate homologous sequences from bacteria, fungi, plants, animals and humans, for expression of the encoded protein (II) in prokaryotic or eukaryotic cells, for inhibiting expression of (II) (when in antisense orientation), and to produce transgenic crop plants. The transgenic plants have modified nucleobase transport properties, e.g. altered affinity and substrate specificity that may result in more efficient nucleobase transport in leaves, changes in apical dominance, flowering behaviour and senescence, or improved distribution of pesticides. This sequence represents the Arabidopsis

Aau57026 Propionib  
Abm53545 Propionib  
Aag28227 Arabidops  
Aag28226 Arabidops  
Aag28225 Arabidops  
Aae22124 Human 543  
Ada20211 Sulphate  
Abb49383 Listeria  
Aag20806 Arabidops  
Abb60159 Drosophil  
Abp25655 Streptoco  
Aab18970 Amino aci  
Abp97205 Tumour-as  
Aam93831 Human pol  
Aam39861 Human pol  
Aag95714 Human pro  
Aam41647 Human pol  
Aag20807 Arabidops  
Aag96821 E. faeciu  
Aar54876 Human H13

26 134.5 7.5 361 4 AAU57026  
27 134.5 7.5 361 6 ABM53545  
28 129 7.2 292 3 AAG28227  
29 129 7.2 390 3 AAG28226  
30 129 7.2 401 3 AAG28225  
31 128.5 7.1 328 5 AAE22124  
32 128.5 7.1 328 5 ADA20211  
33 125.5 7.0 348 5 ABB49383  
34 125 7.0 370 3 AAG20806  
35 124 6.9 352 4 ABB60159  
36 123.5 6.9 674 5 ABP25655  
37 123 6.8 351 3 AAB18970  
38 123 6.8 351 6 ABP97205  
39 123 6.8 364 4 AAM93831  
40 123 6.8 364 4 AAM39861  
41 123 6.8 364 4 AAG95714  
42 123 6.8 370 4 AAM41647  
43 122.5 6.8 365 3 AAG20807  
44 122.5 6.8 376 7 ADC96821  
45 121 6.7 628 2 AAR54876

Mon Sep 27 08:51:15 2004

```

CC thaliana PUP1 protein which is described in the method of the invention
XX
SQ Sequence 356 AA;
Query Match 100.0%; Score 1799; DB 3; Length 356;
Best Local Similarity 100.0%; Pred. No. 5.3e-192;
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;
QY 1 MKNGLIINCIILITGTCGGPILLTRLYFTNGGKRIWFMFSLTAGFPILLPLVLSFLSR 60
DB 1 MKNGLIINCIILITGTCGGPILLTRLYFTNGGKRIWFMFSLTAGFPILLPLVLSFLSR 60
QY 61 RRGNNPNNNAENKRTKFLMETPLFIASIVIGLTLGLDNYLYSGLAYLPVSTSLIIG 120
DB 61 RRGNNPNNNAENKRTKFLMETPLFIASIVIGLTLGLDNYLYSGLAYLPVSTSLIIG 120
QY 121 TQAFNALFAFLLVKQKFTPFPSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
DB 121 TQAFNALFAFLLVKQKFTPFPSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
QY 181 VVAALLYAFILPLVELTYKKARQETFFPLVLEIQMWMLAATFFCVIGMFVGDPKVIAR 240
DB 181 VVAALLYAFILPLVELTYKKARQETFFPLVLEIQMWMLAATFFCVIGMFVGDPKVIAR 240
QY 241 EAREFKIGGSVFYALIVITGIIWQGFGLGAGIVFCASSLASGVLISVLLPVTEVFAVV 300
DB 241 EAREFKIGGSVFYALIVITGIIWQGFGLGAGIVFCASSLASGVLISVLLPVTEVFAVV 300
QY 301 CFREKFOAEKGVSLLLSLMGFVSFYGFEPKSGKKVVDKPPPETELPILPVSDYVA 356
DB 301 CFREKFOAEKGVSLLLSLMGFVSFYGFEPKSGKKVVDKPPPETELPILPVSDYVA 356

RESULT 2
ABB91157
ID ABB91157 standard; protein; 356 AA.
XX
AC ABB91157;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 368.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
PS Claim 5; SEQ ID NO 368; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC nucleic acid or amino acid sequences from plant with nucleic acid or
CC amino acid sequences from non-plant organisms using suitable search
CC parameters, where plant sequences having an E-value greater by a factor
CC

```

```

CC of 3 than the E-value of most similar non-plant sequences are selected.
CC The polypeptides or nucleic acids encoding them are useful for
CC identifying modulators. The identified modulators are useful as
CC herbicides
XX
SQ Sequence 356 AA;
Query Match 99.7%; Score 1792; DB 5; Length 356;
Best Local Similarity 99.7%; Pred. No. 2.5e-191;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;
QY 1 MKNGLIINCIILITGTCGGPILLTRLYFTNGGKRIWFMFSLTAGFPILLPLVLSFLSR 60
DB 1 MKNGLIINCIILITGTCGGPILLTRLYFTNGGKRIWFMFSLTAGFPILLPLVLSFLSR 60
QY 61 RRGNNPNNNAENKRTKFLMETPLFIASIVIGLTLGLDNYLYSGLAYLPVSTSLIIG 120
DB 61 RRGNNPNNNAENKRTKFLMETPLFIASIVIGLTLGLDNYLYSGLAYLPVSTSLIIG 120
QY 121 TQAFNALFAFLLVKQKFTPFPSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
DB 121 TQAFNALFAFLLVKQKFTPFPSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
QY 181 VVAALLYAFILPLVELTYKKARQETFFPLVLEIQMWMLAATFFCVIGMFVGDPKVIAR 240
DB 181 VVAALLYAFILPLVELTYKKARQETFFPLVLEIQMWMLAATFFCVIGMFVGDPKVIAR 240
QY 241 EAREFKIGGSVFYALIVITGIIWQGFGLGAGIVFCASSLASGVLISVLLPVTEVFAVV 300
DB 241 EAREFKIGGSVFYALIVITGIIWQGFGLGAGIVFCASSLASGVLISVLLPVTEVFAVV 300
QY 301 CFREKFOAEKGVSLLLSLMGFVSFYGFEPKSGKKVVDKPPPETELPILPVSDYVA 356
DB 301 CFREKFOAEKGVSLLLSLMGFVSFYGFEPKSGKKVVDKPPPETELPILPVSDYVA 356

RESULT 3
ABB91156
ID ABB91156 standard; protein; 351 AA.
XX
AC ABB91156;
XX
DT 31-MAY-2002 (first entry)
XX
DE Herbicidally active polypeptide SEQ ID NO 367.
XX
KW Herbicidal; plant; agriculture; herbicide.
XX
OS Arabidopsis thaliana.
XX
PN WO200210210-A2.
XX
PD 07-FEB-2002.
XX
PF 28-AUG-2001; 2001WO-EP009892.
XX
PR 28-AUG-2001; 2001WO-EP009892.
XX
PA (FARB ) BAYER AG.
XX
PI Tietjen K, Weidler M;
XX
DR WPI; 2002-269010/31.
XX
PT Identifying plant target proteins for herbicidally active compounds,
PT comprising aligning and comparing nucleic acid or amino acid sequences
PT from plant with nucleic acid or amino acid sequences from non-plant
PT organisms.
XX
PS Claim 5; SEQ ID NO 367; 261pp + Sequence Listing; English.
XX
CC The invention relates to identifying target proteins (ABB90790-ABB94016)
CC for herbicidally active compounds, comprising aligning and comparing
CC

```

23-APR-1999,	PR	99US-0130510P
23-APR-1999,	PR	99US-0130891P
28-APR-1999,	PR	99US-0131449P
30-APR-1999,	PR	99US-0132048P
30-APR-1999,	PR	99US-0132407P
04-MAY-1999,	PR	99US-0132484P
05-MAY-1999,	PR	99US-0132485P
06-MAY-1999,	PR	99US-0132486P
06-MAY-1999,	PR	99US-0132487P
07-MAY-1999,	PR	99US-0132863P
11-MAY-1999,	PR	99US-0134256P
14-MAY-1999,	PR	99US-0134218P
14-MAY-1999,	PR	99US-0134219P
14-MAY-1999,	PR	99US-0134221P
14-MAY-1999,	PR	99US-0134370P
18-MAY-1999,	PR	99US-0134768P
19-MAY-1999,	PR	99US-0134941P
20-MAY-1999,	PR	99US-0135124P
21-MAY-1999,	PR	99US-0135353P
24-MAY-1999,	PR	99US-0135529P
25-MAY-1999,	PR	99US-0136021P
27-MAY-1999,	PR	99US-0136392P
28-MAY-1999,	PR	99US-0136782P
01-JUN-1999,	PR	99US-0137222P
03-JUN-1999,	PR	99US-0137528P
07-JUN-1999,	PR	99US-0137502P
08-JUN-1999,	PR	99US-0137724P
10-JUN-1999,	PR	99US-0138094P
10-JUN-1999,	PR	99US-0138540P
10-JUN-1999,	PR	99US-0138847P
16-JUN-1999,	PR	99US-0139119P
16-JUN-1999,	PR	99US-0139452P
18-JUN-1999,	PR	99US-0139457P
18-JUN-1999,	PR	99US-0139458P
18-JUN-1999,	PR	99US-0139459P
18-JUN-1999,	PR	99US-0139460P
18-JUN-1999,	PR	99US-0139461P
18-JUN-1999,	PR	99US-0139462P
18-JUN-1999,	PR	99US-0139463P
18-JUN-1999,	PR	99US-0139750P
18-JUN-1999,	PR	99US-0139763P
21-JUN-1999,	PR	99US-0139817P
22-JUN-1999,	PR	99US-0139899P
23-JUN-1999,	PR	99US-0140353P
23-JUN-1999,	PR	99US-0140354P
24-JUN-1999,	PR	99US-0140695P
28-JUN-1999,	PR	99US-0140823P
28-JUN-1999,	PR	99US-0140991P
30-JUN-1999,	PR	99US-0141287P
01-JUL-1999,	PR	99US-0141842P
01-JUL-1999,	PR	99US-0142154P
02-JUL-1999,	PR	99US-0142055P
06-JUL-1999,	PR	99US-0142390P
08-JUL-1999,	PR	99US-0142803P
09-JUL-1999,	PR	99US-0142920P
12-JUL-1999,	PR	99US-0142977P
13-JUL-1999,	PR	99US-0143542P
14-JUL-1999,	PR	99US-0143624P
15-JUL-1999,	PR	99US-0144005P
16-JUL-1999,	PR	99US-0144085P
16-JUL-1999,	PR	99US-0144086P
19-JUL-1999,	PR	99US-0144325P
19-JUL-1999,	PR	99US-0144331P
19-JUL-1999,	PR	99US-0144332P
19-JUL-1999,	PR	99US-0144333P
19-JUL-1999,	PR	99US-0144335P
20-JUL-1999,	PR	99US-0144352P

PR 20-JUL-1999;	99US-0144632P.
PR 20-JUL-1999;	99US-0144884P.
PR 21-JUL-1999;	99US-0144814P.
PR 21-JUL-1999;	99US-0145086P.
PR 21-JUL-1999;	99US-0145088P.
PR 22-JUL-1999;	99US-0145088P.
PR 22-JUL-1999;	99US-0145087P.
PR 22-JUL-1999;	99US-0145089P.
PR 22-JUL-1999;	99US-0145192P.
PR 23-JUL-1999;	99US-0145145P.
PR 23-JUL-1999;	99US-0145218P.
PR 23-JUL-1999;	99US-0145224P.
PR 26-JUL-1999;	99US-0145276P.
PR 27-JUL-1999;	99US-0145913P.
PR 27-JUL-1999;	99US-0145918P.
PR 27-JUL-1999;	99US-0145919P.
PR 28-JUL-1999;	99US-0145951P.
PR 02-AUG-1999;	99US-0146388P.
PR 02-AUG-1999;	99US-0146388P.
PR 03-AUG-1999;	99US-0146389P.
PR 03-AUG-1999;	99US-0147038P.
PR 04-AUG-1999;	99US-0147204P.
PR 05-AUG-1999;	99US-0147192P.
PR 06-AUG-1999;	99US-0147260P.
PR 06-AUG-1999;	99US-0147303P.
PR 09-AUG-1999;	99US-0147416P.
PR 09-AUG-1999;	99US-0147493P.
PR 09-AUG-1999;	99US-0147935P.
PR 10-AUG-1999;	99US-0148171P.
PR 11-AUG-1999;	99US-0148319P.
PR 12-AUG-1999;	99US-0148341P.
PR 13-AUG-1999;	99US-0148565P.
PR 13-AUG-1999;	99US-0148684P.
PR 16-AUG-1999;	99US-0149368P.
PR 17-AUG-1999;	99US-0149175P.
PR 18-AUG-1999;	99US-0149426P.
PR 20-AUG-1999;	99US-0149722P.
PR 20-AUG-1999;	99US-0149723P.
PR 20-AUG-1999;	99US-0149929P.
PR 23-AUG-1999;	99US-0149902P.
PR 23-AUG-1999;	99US-0149930P.
PR 25-AUG-1999;	99US-0150566P.
PR 26-AUG-1999;	99US-0150884P.
PR 27-AUG-1999;	99US-0151065P.
PR 27-AUG-1999;	99US-0151066P.
PR 27-AUG-1999;	99US-0151080P.
PR 30-AUG-1999;	99US-0151303P.
PR 31-AUG-1999;	99US-0151438P.
PR 01-SEP-1999;	99US-0151930P.
PR 07-SEP-1999;	99US-0152363P.
PR 10-SEP-1999;	99US-0153070P.
PR 13-SEP-1999;	99US-0153758P.
PR 15-SEP-1999;	99US-0154018P.
PR 16-SEP-1999;	99US-0154039P.
PR 20-SEP-1999;	99US-0154779P.
PR 22-SEP-1999;	99US-0155139P.
PR 23-SEP-1999;	99US-0155486P.
PR 24-SEP-1999;	99US-0155659P.
PR 28-SEP-1999;	99US-0156458P.
PR 28-SEP-1999;	99US-0156596P.
PR 08-OCT-1999;	99US-0157117P.
PR 08-OCT-1999;	99US-0157753P.
PR 08-OCT-1999;	99US-0157865P.
PR 08-OCT-1999;	99US-0158029P.
PR 08-OCT-1999;	99US-0158232P.
PR 13-OCT-1999;	99US-0158369P.
PR 13-OCT-1999;	99US-0159293P.
PR 13-OCT-1999;	99US-0159294P.
PR 13-OCT-1999;	99US-0159329P.
PR 14-OCT-1999;	99US-0159330P.
PR 14-OCT-1999;	99US-0159331P.

PR 14-OCT-1999;	99US-0159637P.
PR 14-OCT-1999;	99US-0159638P.
PR 18-OCT-1999;	99US-0159584P.
PR 21-OCT-1999;	99US-0160741P.
PR 21-OCT-1999;	99US-0160767P.
PR 21-OCT-1999;	99US-0160768P.
PR 21-OCT-1999;	99US-0160770P.
PR 21-OCT-1999;	99US-0160814P.
PR 21-OCT-1999;	99US-0160815P.
PR 22-OCT-1999;	99US-0160980P.
PR 22-OCT-1999;	99US-0160981P.
PR 22-OCT-1999;	99US-0160989P.
PR 25-OCT-1999;	99US-0161404P.
PR 25-OCT-1999;	99US-0161405P.
PR 25-OCT-1999;	99US-0161406P.
PR 26-OCT-1999;	99US-0161359P.
PR 26-OCT-1999;	99US-0161360P.
PR 26-OCT-1999;	99US-0161361P.
PR 26-OCT-1999;	99US-0161920P.
PR 28-OCT-1999;	99US-0161922P.
PR 28-OCT-1999;	99US-0161933P.
PR 29-OCT-1999;	99US-0162142P.

Query Match

Best Local Similarity

Matches 177; Conservative

49.6%; 99.4%; 0; Mismatches 1; Indels 0; Gaps 0;

Score 892; DB 3; Length 178;

Qy	179	MTVVALLYAFILPLVELTYKKARQETTFPLVLEIQMVLMCLAAATFFCVIGMFIVGDFKVI	238
Db	1	MTVVALLYAFILPLVELTYKKARQETTFPLVLEIQMVLMCLAAATFFCVIGMFIVGDFKVI	60
Qy	239	AREAREFKIGSVFYFVYALIVITGIWQFGLGIVFCASSLASGVLISVLLPVTVEFA	298
Db	61	AXEAREFKIGSVFYFVYALIVITGIWQFGLGIVFCASSLASGVLISVLLPVTVEFA	120
Qy	299	VVCFREKFOAEKGVSLLSLWGFVSFYFGFKSGKGVVDPQPPETELPILPVSDYVA	356
Db	121	VVCFREKFOAEKGVSLLSLWGFVSFYFGFKSGKGVVDPQPPETELPILPVSDYVA	178

RESULT 5

AAG15103

ID AAG15103 standard; protein; 144 AA.

XX AAG15103;

AC AAG15103;

XX

DT 17-OCT-2000 (first entry)

XX

DE Arabidopsis thaliana protein fragment SEQ ID NO: 15224.

XX

KW Protein identification; signal transduction pathway; metabolic pathway;

KW hybridisation assay; genetic mapping; gene expression control; promoter;

KW termination sequence.

XX

OS Arabidopsis thaliana.

XX

XX EPI033405-A2.

PN

PD 06-SEP-2000.

XX

PF 25-FEB-2000; 2000EP-00301439.

XX

PR 25-FEB-1999; 99US-0121825P.

PR 05-MAR-1999; 99US-0123180P.

PR 09-MAR-1999; 99US-0123548P.

PR 23-MAR-1999; 99US-0125788P.

PR 25-MAR-1999; 99US-0126264P.

PR 29-MAR-1999; 99US-0126785P.

PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.

PR 08-APR-1999; 99US-0128714P.

PR 16-APR-1999; 99US-0129845P.

PR 19-APR-1999; 99US-0130077P.

PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132484P.  
PR 05-MAY-1999; 99US-0132485P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 11-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-0134256P.  
PR 14-MAY-1999; 99US-0134218P.  
PR 14-MAY-1999; 99US-0134219P.  
PR 14-MAY-1999; 99US-0134221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 24-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 26-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 23-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 20-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-015659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.

PR	14-OCT-1999;	99US-0159331P.	PR	19-APR-1999;	99US-0130077P.
PR	14-OCT-1999;	99US-0159637P.	PR	21-APR-1999;	99US-0130449P.
PR	14-OCT-1999;	99US-0159638P.	PR	23-APR-1999;	99US-0130510P.
PR	18-OCT-1999;	99US-0159584P.	PR	23-APR-1999;	99US-0130891P.
PR	21-OCT-1999;	99US-0160741P.	PR	28-APR-1999;	99US-0131449P.
PR	21-OCT-1999;	99US-0160767P.	PR	30-APR-1999;	99US-0132048P.
PR	21-OCT-1999;	99US-0160768P.	PR	30-APR-1999;	99US-0132407P.
PR	21-OCT-1999;	99US-0160770P.	PR	04-MAY-1999;	99US-0132484P.
PR	21-OCT-1999;	99US-0160814P.	PR	05-MAY-1999;	99US-0132485P.
PR	21-OCT-1999;	99US-0160815P.	PR	06-MAY-1999;	99US-0132486P.
PR	21-OCT-1999;	99US-0160980P.	PR	07-MAY-1999;	99US-0132487P.
PR	22-OCT-1999;	99US-0160981P.	PR	11-MAY-1999;	99US-0132566P.
PR	22-OCT-1999;	99US-0160989P.	PR	14-MAY-1999;	99US-0134218P.
PR	25-OCT-1999;	99US-0161404P.	PR	14-MAY-1999;	99US-0134221P.
PR	25-OCT-1999;	99US-0161405P.	PR	14-MAY-1999;	99US-0134370P.
PR	25-OCT-1999;	99US-0161406P.	PR	18-MAY-1999;	99US-0134768P.
PR	26-OCT-1999;	99US-0161359P.	PR	19-MAY-1999;	99US-0134941P.
PR	26-OCT-1999;	99US-0161360P.	PR	20-MAY-1999;	99US-0135124P.
PR	26-OCT-1999;	99US-0161361P.	PR	21-MAY-1999;	99US-0135553P.
PR	28-OCT-1999;	99US-0161920P.	PR	24-MAY-1999;	99US-0135629P.
PR	28-OCT-1999;	99US-0161992P.	PR	25-MAY-1999;	99US-0136021P.
PR	28-OCT-1999;	99US-0161993P.	PR	27-MAY-1999;	99US-0136392P.
PR	29-OCT-1999;	99US-0162142P.	PR	28-MAY-1999;	99US-0136782P.
Query Match 40.4%; Score 727; DB 3; Length 144;					99US-0137222P.
Best Local Similarity 99.3%; Pred. No. 9.5e-73;					99US-0137528P.
Matches 143; Conservative 0; Mismatches 1; Indels 0; Gaps 0;					99US-0137528P.
QY	1	MKNGLIINCIILTTGTGCGPLLTPLRYFTNGGKRIWFMSFLSTAGFPPIILPLLVSLR 60	PR	03-JUN-1999;	99US-0137528P.
DB	1	MKNGLIINCIILTTGTGCGPLLTPLRYFTNGGKRIWFMSFLSTAGFPPIILPLLVSLR 60	PR	04-JUN-1999;	99US-0137528P.
QY	61	RRGNRPNNNAENKRTKLFMETPLFIASIVIGLTLGLDNLVSYGLAYLPVSTSSLIIG 120	PR	07-JUN-1999;	99US-0137724P.
DB	61	RRSNRPNNNAENKRTKLFMETPLFIASIVIGLTLGLDNLVSYGLAYLPVSTSSLIIG 120	PR	10-JUN-1999;	99US-0138094P.
QY	121	TQAFNALPAPLLVKKQFTPPFSIN 144	PR	10-JUN-1999;	99US-0138540P.
DB	121	TQAFNALPAPLLVKKQFTPPFSIN 144	PR	14-JUN-1999;	99US-0138847P.
RESULT 6					99US-0139119P.
AAGI5102					99US-0139452P.
ID	AAGI5102 standard; protein; 164 AA.				99US-0139453P.
XX					99US-0139492P.
XX	AAGI5102;				99US-0139454P.
XX	AC				99US-0139455P.
XX	17-OCT-2000 (first entry)				99US-0139456P.
XX					99US-0139457P.
DE	Arabidopsis thaliana protein fragment SEQ ID NO: 15223.				99US-0139458P.
XX					99US-0139459P.
KW	Protein identification; signal transduction pathway; metabolic pathway;				99US-0139460P.
KW	hybridisation assay; Genetic mapping; gene expression control; promoter;				99US-0139461P.
XX	termination sequence.				99US-0139462P.
XX					99US-0139463P.
OS	Arabidopsis thaliana.				99US-0139750P.
XX					99US-0139750P.
PN	EP1033405-A2.				99US-0139763P.
XX					99US-0139817P.
PD	06-SEP-2000.				99US-0139899P.
XX					99US-0140353P.
PF	25-FEB-2000; 2000EP-00301439.				99US-0140354P.
XX					99US-0140695P.
PR	25-FEB-1999;				99US-0140823P.
PR	05-MAR-1999;				99US-0140991P.
PR	09-MAR-1999;				99US-0141287P.
PR	23-MAR-1999;				99US-0141842P.
PR	23-MAR-1999;				99US-0142154P.
PR	29-MAR-1999;				99US-0142055P.
PR	01-APR-1999;				99US-0142390P.
PR	06-APR-1999;				99US-0142803P.
PR	08-APR-1999;				99US-0142977P.
PR	16-APR-1999;				99US-0142977P.
PR					99US-0143542P.
PR					99US-0143624P.
PR					99US-0144005P.
PR					99US-0144085P.
PR					99US-0144086P.
PR					99US-0144325P.
PR					99US-0144331P.
PR					99US-0144332P.
PR					99US-0144333P.
PR					99US-0144334P.

```
PR 19-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144335P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144632P.
PR 20-JUL-1999; 99US-0144884P.
PR 21-JUL-1999; 99US-0144814P.
PR 21-JUL-1999; 99US-0145086P.
PR 21-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145088P.
PR 22-JUL-1999; 99US-0145087P.
PR 22-JUL-1999; 99US-0145089P.
PR 22-JUL-1999; 99US-0145192P.
PR 23-JUL-1999; 99US-0145145P.
PR 23-JUL-1999; 99US-0145218P.
PR 23-JUL-1999; 99US-0145224P.
PR 26-JUL-1999; 99US-0145276P.
PR 27-JUL-1999; 99US-0145913P.
PR 27-JUL-1999; 99US-0145918P.
PR 27-JUL-1999; 99US-0145919P.
PR 28-JUL-1999; 99US-0145951P.
PR 02-AUG-1999; 99US-0146386P.
PR 02-AUG-1999; 99US-0146388P.
PR 02-AUG-1999; 99US-0146388P.
PR 03-AUG-1999; 99US-0147038P.
PR 04-AUG-1999; 99US-0147204P.
PR 04-AUG-1999; 99US-0147302P.
PR 05-AUG-1999; 99US-0147192P.
PR 05-AUG-1999; 99US-0147260P.
PR 06-AUG-1999; 99US-0147303P.
PR 06-AUG-1999; 99US-0147416P.
PR 09-AUG-1999; 99US-0147493P.
PR 09-AUG-1999; 99US-0147935P.
PR 10-AUG-1999; 99US-0148171P.
PR 11-AUG-1999; 99US-0148319P.
PR 12-AUG-1999; 99US-0148341P.
PR 13-AUG-1999; 99US-0148565P.
PR 13-AUG-1999; 99US-0148684P.
PR 16-AUG-1999; 99US-0149368P.
PR 17-AUG-1999; 99US-0149175P.
PR 18-AUG-1999; 99US-0149426P.
PR 20-AUG-1999; 99US-0149722P.
PR 20-AUG-1999; 99US-0149723P.
PR 20-AUG-1999; 99US-0149929P.
PR 23-AUG-1999; 99US-0149902P.
PR 23-AUG-1999; 99US-0149930P.
PR 25-AUG-1999; 99US-0150566P.
PR 26-AUG-1999; 99US-0150884P.
PR 27-AUG-1999; 99US-0151065P.
PR 27-AUG-1999; 99US-0151066P.
PR 27-AUG-1999; 99US-0151080P.
PR 30-AUG-1999; 99US-0151303P.
PR 31-AUG-1999; 99US-0151438P.
PR 01-SEP-1999; 99US-0151930P.
PR 07-SEP-1999; 99US-0152363P.
PR 10-SEP-1999; 99US-0153070P.
PR 13-SEP-1999; 99US-0153758P.
PR 15-SEP-1999; 99US-0154018P.
PR 16-SEP-1999; 99US-0154039P.
PR 20-SEP-1999; 99US-0154779P.
PR 22-SEP-1999; 99US-0155139P.
PR 23-SEP-1999; 99US-0155486P.
PR 24-SEP-1999; 99US-0155659P.
PR 28-SEP-1999; 99US-0156458P.
PR 29-SEP-1999; 99US-0156596P.
PR 04-OCT-1999; 99US-0157117P.
PR 05-OCT-1999; 99US-0157753P.
PR 06-OCT-1999; 99US-0157865P.
PR 07-OCT-1999; 99US-0158029P.
PR 08-OCT-1999; 99US-0158232P.
PR 12-OCT-1999; 99US-0158369P.
PR 13-OCT-1999; 99US-0159293P.
PR 13-OCT-1999; 99US-0159294P.
PR 13-OCT-1999; 99US-0159295P.
PR 14-OCT-1999; 99US-0159329P.

PR 14-OCT-1999; 99US-0159330P.
PR 14-OCT-1999; 99US-0159331P.
PR 14-OCT-1999; 99US-0159637P.
PR 14-OCT-1999; 99US-0159638P.
PR 18-OCT-1999; 99US-0159584P.
PR 21-OCT-1999; 99US-0160741P.
PR 21-OCT-1999; 99US-0160767P.
PR 21-OCT-1999; 99US-0160768P.
PR 21-OCT-1999; 99US-0160770P.
PR 21-OCT-1999; 99US-0160814P.
PR 21-OCT-1999; 99US-0160815P.
PR 22-OCT-1999; 99US-0160980P.
PR 22-OCT-1999; 99US-0160981P.
PR 22-OCT-1999; 99US-0160989P.
PR 25-OCT-1999; 99US-0161404P.
PR 25-OCT-1999; 99US-0161405P.
PR 25-OCT-1999; 99US-0161406P.
PR 26-OCT-1999; 99US-0161359P.
PR 26-OCT-1999; 99US-0161360P.
PR 26-OCT-1999; 99US-0161361P.
PR 28-OCT-1999; 99US-0161920P.
PR 28-OCT-1999; 99US-0161992P.
PR 28-OCT-1999; 99US-0161993P.
PR 23-OCT-1999; 99US-0162142P.

Query Match 40.4%; Score 727; DB 3; Length 164;
Best Local Similarity 99.3%; Pred. No. 1.le-72;
Matches 143; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKNGLIINCIILITICTCGPLLTRLYFTNGGKRIWMFSLTAGPPIILIPLLVPSLSR 60
Db 21 MKNGLIINCIILITICTCGPLLTRLYFTNGGKRIWMFSLTAGPPIILIPLLVPSLSR 80
QY 61 RRGNRPNNAENKRTKLFMETPLFIASIVIGLLTGLDNLVSYGLAYLPVSTSLIIG 120
Db 81 RRGNRPNNAENKRTKLFMETPLFIASIVIGLLTGLDNLVSYGLAYLPVSTSLIIG 140
QY 121 TQLAFNALFAFLVVKQKTFPFSIN 144
Db 141 TQLAFNALFAFLVVKQKTFPFSIN 164

RESULT 7
AAG22805
ID AAG22805 standard; protein; 367 AA.
XX
AC AAG22805;
XX
DT 17-OCT-2000 (first entry)
XX
DE Arabidopsis thaliana protein fragment SEQ ID NO: 25874.
XX
KW Protein identification; signal transduction pathway; metabolic pathway;
KW hybridisation assay; genetic mapping; gene expression control; promoter;
XX termination sequence.
OS Arabidopsis thaliana.
XX
PN EP1033405-A2.
XX
PD 06-SEP-2000.
XX
PF 25-FEB-2000; 2000EP-00301439.
XX
PR 25-FEB-1999; 99US-0121825P.
PR 05-MAR-1999; 99US-0123180P.
PR 09-MAR-1999; 99US-0123548P.
PR 23-MAR-1999; 99US-0125788P.
PR 25-MAR-1999; 99US-0126264P.
PR 29-MAR-1999; 99US-0126785P.
PR 01-APR-1999; 99US-0127462P.
PR 06-APR-1999; 99US-0128234P.
PR 08-APR-1999; 99US-0128714P.
```

us-09-913-767-8.rag

Mon Sep 27 08:51:15 2004

PR 16-APR-1999;	99US-0129845P.	PR 19-JUL-1999;	99US-0144334P.
PR 19-APR-1999;	99US-0130077P.	PR 19-JUL-1999;	99US-0144335P.
PR 21-APR-1999;	99US-0130449P.	PR 20-JUL-1999;	99US-0144352P.
PR 23-APR-1999;	99US-0130510P.	PR 20-JUL-1999;	99US-0144632P.
PR 23-APR-1999;	99US-0130891P.	PR 21-JUL-1999;	99US-0144884P.
PR 28-APR-1999;	99US-0131449P.	PR 21-JUL-1999;	99US-0144814P.
PR 30-APR-1999;	99US-0132048P.	PR 21-JUL-1999;	99US-0145086P.
PR 30-APR-1999;	99US-0132407P.	PR 21-JUL-1999;	99US-0145088P.
PR 04-MAY-1999;	99US-0132484P.	PR 22-JUL-1999;	99US-0145085P.
PR 05-MAY-1999;	99US-0132485P.	PR 22-JUL-1999;	99US-0145087P.
PR 06-MAY-1999;	99US-0132486P.	PR 22-JUL-1999;	99US-0145089P.
PR 06-MAY-1999;	99US-0132487P.	PR 22-JUL-1999;	99US-0145192P.
PR 07-MAY-1999;	99US-0132863P.	PR 23-JUL-1999;	99US-0145145P.
PR 11-MAY-1999;	99US-0134256P.	PR 23-JUL-1999;	99US-0145218P.
PR 14-MAY-1999;	99US-0134218P.	PR 23-JUL-1999;	99US-0145224P.
PR 14-MAY-1999;	99US-0134219P.	PR 26-JUL-1999;	99US-0145276P.
PR 14-MAY-1999;	99US-0134221P.	PR 27-JUL-1999;	99US-0145913P.
PR 14-MAY-1999;	99US-0134370P.	PR 27-JUL-1999;	99US-0145918P.
PR 18-MAY-1999;	99US-0134768P.	PR 28-JUL-1999;	99US-0145919P.
PR 19-MAY-1999;	99US-0134941P.	PR 02-AUG-1999;	99US-0145951P.
PR 20-MAY-1999;	99US-0135124P.	PR 02-AUG-1999;	99US-0146386P.
PR 21-MAY-1999;	99US-0135353P.	PR 02-AUG-1999;	99US-0146388P.
PR 25-MAY-1999;	99US-0135629P.	PR 03-AUG-1999;	99US-0146389P.
PR 27-MAY-1999;	99US-0136021P.	PR 03-AUG-1999;	99US-0147038P.
PR 28-MAY-1999;	99US-0136392P.	PR 04-AUG-1999;	99US-0147204P.
PR 01-JUN-1999;	99US-0136782P.	PR 04-AUG-1999;	99US-0147302P.
PR 03-JUN-1999;	99US-0137222P.	PR 05-AUG-1999;	99US-0147192P.
PR 04-JUN-1999;	99US-0137528P.	PR 05-AUG-1999;	99US-0147260P.
PR 07-JUN-1999;	99US-0137502P.	PR 06-AUG-1999;	99US-0147303P.
PR 08-JUN-1999;	99US-0137724P.	PR 06-AUG-1999;	99US-0147416P.
PR 08-JUN-1999;	99US-0138094P.	PR 09-AUG-1999;	99US-0147493P.
PR 10-JUN-1999;	99US-0138540P.	PR 09-AUG-1999;	99US-0147935P.
PR 10-JUN-1999;	99US-0138847P.	PR 10-AUG-1999;	99US-0148171P.
PR 14-JUN-1999;	99US-0139119P.	PR 11-AUG-1999;	99US-0148319P.
PR 16-JUN-1999;	99US-0139452P.	PR 12-AUG-1999;	99US-0148341P.
PR 16-JUN-1999;	99US-0139453P.	PR 13-AUG-1999;	99US-0148565P.
PR 17-JUN-1999;	99US-0139492P.	PR 13-AUG-1999;	99US-0148684P.
PR 17-JUN-1999;	99US-0139454P.	PR 16-AUG-1999;	99US-0149368P.
PR 18-JUN-1999;	99US-0139455P.	PR 17-AUG-1999;	99US-0149175P.
PR 18-JUN-1999;	99US-0139456P.	PR 18-AUG-1999;	99US-0149426P.
PR 18-JUN-1999;	99US-0139457P.	PR 20-AUG-1999;	99US-0149722P.
PR 18-JUN-1999;	99US-0139458P.	PR 20-AUG-1999;	99US-0149723P.
PR 18-JUN-1999;	99US-0139459P.	PR 20-AUG-1999;	99US-0149929P.
PR 18-JUN-1999;	99US-0139460P.	PR 23-AUG-1999;	99US-0149902P.
PR 18-JUN-1999;	99US-0139461P.	PR 25-AUG-1999;	99US-0150566P.
PR 18-JUN-1999;	99US-0139462P.	PR 26-AUG-1999;	99US-0150884P.
PR 18-JUN-1999;	99US-0139463P.	PR 27-AUG-1999;	99US-0151065P.
PR 18-JUN-1999;	99US-0139750P.	PR 27-AUG-1999;	99US-0151066P.
PR 18-JUN-1999;	99US-0139763P.	PR 30-AUG-1999;	99US-0151080P.
PR 21-JUN-1999;	99US-0139817P.	PR 31-AUG-1999;	99US-0151303P.
PR 22-JUN-1999;	99US-0139899P.	PR 01-SEP-1999;	99US-0151930P.
PR 23-JUN-1999;	99US-0140353P.	PR 01-SEP-1999;	99US-0152363P.
PR 23-JUN-1999;	99US-0140354P.	PR 10-SEP-1999;	99US-0153758P.
PR 24-JUN-1999;	99US-0140695P.	PR 13-SEP-1999;	99US-0154018P.
PR 28-JUN-1999;	99US-0140823P.	PR 15-SEP-1999;	99US-0154039P.
PR 28-JUN-1999;	99US-0140991P.	PR 16-SEP-1999;	99US-0154779P.
PR 30-JUN-1999;	99US-0141287P.	PR 20-SEP-1999;	99US-0155139P.
PR 01-JUL-1999;	99US-0141842P.	PR 23-SEP-1999;	99US-0155486P.
PR 01-JUL-1999;	99US-0142154P.	PR 24-SEP-1999;	99US-0155659P.
PR 02-JUL-1999;	99US-0142055P.	PR 28-SEP-1999;	99US-0156458P.
PR 06-JUL-1999;	99US-0142390P.	PR 29-SEP-1999;	99US-0156596P.
PR 08-JUL-1999;	99US-0142803P.	PR 04-OCT-1999;	99US-0157117P.
PR 09-JUL-1999;	99US-0142920P.	PR 05-OCT-1999;	99US-0157753P.
PR 12-JUL-1999;	99US-0142977P.	PR 06-OCT-1999;	99US-0157865P.
PR 13-JUL-1999;	99US-0143542P.	PR 07-OCT-1999;	99US-0158029P.
PR 14-JUL-1999;	99US-0143624P.	PR 08-OCT-1999;	99US-0158232P.
PR 15-JUL-1999;	99US-0144005P.	PR 12-OCT-1999;	99US-0158369P.
PR 16-JUL-1999;	99US-0144085P.	PR 13-OCT-1999;	99US-0159293P.
PR 16-JUL-1999;	99US-0144086P.	PR 13-OCT-1999;	99US-0159294P.
PR 19-JUL-1999;	99US-0144332P.	PR 13-OCT-1999;	99US-0159295P.
PR 19-JUL-1999;	99US-0144333P.		



PR	14-OCT-1999;	99US-0159329P.	PD	06-SEP-2000.	
PR	14-OCT-1999;	99US-0159330P.	XX		
PR	14-OCT-1999;	99US-0159331P.	PF	25-FEB-2000;	2000EP-00301439.
PR	14-OCT-1999;	99US-0159332P.	XX		
PR	14-OCT-1999;	99US-0159333P.	XX		
PR	14-OCT-1999;	99US-0159334P.	PR	25-FEB-1999;	99US-0121825P.
PR	14-OCT-1999;	99US-0159335P.	PR	03-MAR-1999;	99US-0123180P.
PR	14-OCT-1999;	99US-0159336P.	PR	09-MAR-1999;	99US-0123548P.
PR	14-OCT-1999;	99US-0159337P.	PR	23-MAR-1999;	99US-0125788P.
PR	14-OCT-1999;	99US-0159338P.	PR	25-MAR-1999;	99US-0126264P.
PR	14-OCT-1999;	99US-0159339P.	PR	29-MAR-1999;	99US-0126785P.
PR	14-OCT-1999;	99US-0159340P.	PR	01-APR-1999;	99US-0127462P.
PR	14-OCT-1999;	99US-0160741P.	PR	06-APR-1999;	99US-0128234P.
PR	14-OCT-1999;	99US-0160767P.	PR	08-APR-1999;	99US-0128714P.
PR	14-OCT-1999;	99US-0160768P.	PR	16-APR-1999;	99US-0129845P.
PR	14-OCT-1999;	99US-0160770P.	PR	19-APR-1999;	99US-0130077P.
PR	14-OCT-1999;	99US-0160815P.	PR	21-APR-1999;	99US-0130449P.
PR	14-OCT-1999;	99US-0160814P.	PR	23-APR-1999;	99US-0130510P.
PR	14-OCT-1999;	99US-0160880P.	PR	28-APR-1999;	99US-0130891P.
PR	14-OCT-1999;	99US-0160981P.	PR	30-APR-1999;	99US-0131449P.
PR	14-OCT-1999;	99US-0161040P.	PR	04-MAY-1999;	99US-0132048P.
PR	14-OCT-1999;	99US-0161405P.	PR	05-MAY-1999;	99US-0132484P.
PR	14-OCT-1999;	99US-0161406P.	PR	06-MAY-1999;	99US-0132485P.
PR	14-OCT-1999;	99US-0161359P.	PR	07-MAY-1999;	99US-0132486P.
PR	14-OCT-1999;	99US-0161360P.	PR	11-MAY-1999;	99US-0132863P.
PR	14-OCT-1999;	99US-0161361P.	PR	14-MAY-1999;	99US-0134218P.
PR	14-OCT-1999;	99US-0161362P.	PR	14-MAY-1999;	99US-0134219P.
PR	14-OCT-1999;	99US-0161363P.	PR	14-MAY-1999;	99US-0134221P.
PR	14-OCT-1999;	99US-0161364P.	PR	18-MAY-1999;	99US-0134370P.
PR	14-OCT-1999;	99US-0161365P.	PR	19-MAY-1999;	99US-0134941P.
PR	14-OCT-1999;	99US-0161366P.	PR	21-MAY-1999;	99US-0135124P.
PR	14-OCT-1999;	99US-0161367P.	PR	21-MAY-1999;	99US-0135353P.
PR	14-OCT-1999;	99US-0161368P.	PR	25-MAY-1999;	99US-0135629P.
PR	14-OCT-1999;	99US-0161369P.	PR	27-MAY-1999;	99US-0136392P.
PR	14-OCT-1999;	99US-0161370P.	PR	28-MAY-1999;	99US-0136782P.
PR	14-OCT-1999;	99US-0161371P.	PR	01-JUN-1999;	99US-0137222P.
PR	14-OCT-1999;	99US-0161372P.	PR	03-JUN-1999;	99US-0137528P.
PR	14-OCT-1999;	99US-0161373P.	PR	07-JUN-1999;	99US-0137724P.
PR	14-OCT-1999;	99US-0161374P.	PR	10-JUN-1999;	99US-0138540P.
PR	14-OCT-1999;	99US-0161375P.	PR	10-JUN-1999;	99US-0138847P.
PR	14-OCT-1999;	99US-0161376P.	PR	14-JUN-1999;	99US-0139119P.
PR	14-OCT-1999;	99US-0161377P.	PR	16-JUN-1999;	99US-0139452P.
PR	14-OCT-1999;	99US-0161378P.	PR	16-JUN-1999;	99US-0139453P.
PR	14-OCT-1999;	99US-0161379P.	PR	17-JUN-1999;	99US-0139492P.
PR	14-OCT-1999;	99US-0161380P.	PR	18-JUN-1999;	99US-0139454P.
PR	14-OCT-1999;	99US-0161381P.	PR	18-JUN-1999;	99US-0139455P.
PR	14-OCT-1999;	99US-0161382P.	PR	18-JUN-1999;	99US-0139456P.
PR	14-OCT-1999;	99US-0161383P.	PR	18-JUN-1999;	99US-0139457P.
PR	14-OCT-1999;	99US-0161384P.	PR	18-JUN-1999;	99US-0139458P.
PR	14-OCT-1999;	99US-0161385P.	PR	18-JUN-1999;	99US-0139459P.
PR	14-OCT-1999;	99US-0161386P.	PR	18-JUN-1999;	99US-0139460P.
PR	14-OCT-1999;	99US-0161387P.	PR	18-JUN-1999;	99US-0139461P.
PR	14-OCT-1999;	99US-0161388P.	PR	18-JUN-1999;	99US-0139462P.
PR	14-OCT-1999;	99US-0161389P.	PR	18-JUN-1999;	99US-0139463P.
PR	14-OCT-1999;	99US-0161390P.	PR	18-JUN-1999;	99US-0139750P.
PR	14-OCT-1999;	99US-0161391P.	PR	21-JUN-1999;	99US-0139763P.
PR	14-OCT-1999;	99US-0161392P.	PR	22-JUN-1999;	99US-0139899P.
PR	14-OCT-1999;	99US-0161393P.	PR	23-JUN-1999;	99US-0140353P.
PR	14-OCT-1999;	99US-0161394P.	PR	23-JUN-1999;	99US-0140354P.
PR	14-OCT-1999;	99US-0161395P.	PR	24-JUN-1999;	99US-0140695P.
PR	14-OCT-1999;	99US-0161396P.	PR	28-JUN-1999;	99US-0140823P.
PR	14-OCT-1999;	99US-0161397P.	PR	29-JUN-1999;	99US-0140991P.
PR	14-OCT-1999;	99US-0161398P.	PR	30-JUN-1999;	99US-0141287P.
PR	14-OCT-1999;	99US-0161399P.	PR	01-JUL-1999;	99US-0141842P.
PR	14-OCT-1999;	99US-0161400P.	PR	01-JUL-1999;	99US-0142154P.
PR	14-OCT-1999;	99US-0161401P.	PR	02-JUL-1999;	99US-0142055P.

QY	5	LIINCIILITGTCGPGPLRLFTNGGKRIWMSFLSTAGFPILLLVLSFLRRGN	64
Db	12	LLIVTFPLFFGSIASSLLAKYFYVGGSRWSTWQSSAGFPLLLIYFPHYVLKTTT	71
QY	65	RNPNAENKRTKFLMETPLFIASIVIGLLTGLDNLVSYGLAYLPVSTSSLIIGTOLA	124
Db	72	RRP-----FTFRLH-----LFSLVIGLVGNFPLFSWGTSLPVSTSSLLSTQLV	121
QY	125	FNALFAPLVKQKFTPFSSINAVLLTVGIGIALHSDGDKPAKESKEYVVGFLMTVAA	184
Db	122	FTILSRIVKQITFSNLNCVLLTSSVLLALDSSKDKPSGLTKTKFYGVSTIGAG	181
QY	185	LLYAFILPLVELTYKARQBITPLVLEIQMNCCLAAATPCFVGMFVDFKVIAREARE	244
Db	182	LLFALYLPVTEKLY--RTVICYAMVMEVQLVMEFAATVPATIGMACGGFKEMVKEANH	238
QY	245	EKIGGSVEYVALIVITGII--WQGFELGAIIVFCASSLASGVLSVLLPVTVEFAVVCFR	303
Db	239	VFTKGFTFTWTFAILANVVTWQLSFAATSGMVLTSGITGGICMTALLANVIGGVVAYG	298
QY	304	EKFQAEKGVSLLSLMGFVSFYFGEFKSGKK	334
Db	299	DVFGVKIVSTVLCIWFSSYTYGMYMKKK	329

Query Match 30.8%; Score 553; DB 3; Length 367;  
 Best Local Similarity 36.3%; Pred. No. 9.7e-53;  
 Matches 120; Conservative 65; Mismatches 132; Indels 14; Gaps 4;

RESULT 8  
 AAG40302  
 ID AAG40302 standard; protein; 367 AA.  
 XX AAG40302;  
 AC AAG40302;  
 XX AAG40302;  
 DT 18-OCT-2000 (first entry)  
 DE Arabidopsis thaliana protein fragment SEQ ID NO: 49987.  
 DE DE  
 XX DE  
 XX DE  
 KW Protein identification; signal transduction pathway; metabolic pathway;  
 KW hybridisation assay; genetic mapping; gene expression control; promoter;  
 KW termination sequence.  
 XX  
 OS Arabidopsis thaliana.  
 XX  
 XX  
 PN EP1033405-A2.  
 XX

PR	06-JUL-1999;	99US-0142330P.
PR	08-JUL-1999;	99US-0142803P.
PR	09-JUL-1999;	99US-0142920P.
PR	12-JUL-1999;	99US-0142977P.
PR	13-JUL-1999;	99US-0143542P.
PR	14-JUL-1999;	99US-0143624P.
PR	15-JUL-1999;	99US-0144005P.
PR	16-JUL-1999;	99US-0144085P.
PR	16-JUL-1999;	99US-0144086P.
PR	16-JUL-1999;	99US-0144325P.
PR	19-JUL-1999;	99US-0144331P.
PR	19-JUL-1999;	99US-0144332P.
PR	19-JUL-1999;	99US-0144333P.
PR	19-JUL-1999;	99US-0144334P.
PR	19-JUL-1999;	99US-0144335P.
PR	20-JUL-1999;	99US-0144352P.
PR	20-JUL-1999;	99US-0144632P.
PR	21-JUL-1999;	99US-0144884P.
PR	21-JUL-1999;	99US-0145086P.
PR	21-JUL-1999;	99US-0145088P.
PR	21-JUL-1999;	99US-0145085P.
PR	22-JUL-1999;	99US-0145087P.
PR	22-JUL-1999;	99US-0145089P.
PR	22-JUL-1999;	99US-0145192P.
PR	23-JUL-1999;	99US-0145145P.
PR	23-JUL-1999;	99US-0145218P.
PR	23-JUL-1999;	99US-0145224P.
PR	26-JUL-1999;	99US-0145276P.
PR	26-JUL-1999;	99US-0145913P.
PR	27-JUL-1999;	99US-0145918P.
PR	27-JUL-1999;	99US-0145951P.
PR	28-JUL-1999;	99US-0145951P.
PR	02-AUG-1999;	99US-0146386P.
PR	02-AUG-1999;	99US-0146388P.
PR	02-AUG-1999;	99US-0146389P.
PR	03-AUG-1999;	99US-0147038P.
PR	04-AUG-1999;	99US-0147204P.
PR	04-AUG-1999;	99US-0147302P.
PR	05-AUG-1999;	99US-0147192P.
PR	05-AUG-1999;	99US-0147260P.
PR	06-AUG-1999;	99US-0147303P.
PR	06-AUG-1999;	99US-0147416P.
PR	09-AUG-1999;	99US-0147493P.
PR	09-AUG-1999;	99US-0147935P.
PR	10-AUG-1999;	99US-0148171P.
PR	11-AUG-1999;	99US-0148319P.
PR	12-AUG-1999;	99US-0148341P.
PR	13-AUG-1999;	99US-0148565P.
PR	13-AUG-1999;	99US-0148684P.
PR	16-AUG-1999;	99US-0149368P.
PR	17-AUG-1999;	99US-0149175P.
PR	18-AUG-1999;	99US-0149426P.
PR	20-AUG-1999;	99US-0149722P.
PR	20-AUG-1999;	99US-0149723P.
PR	20-AUG-1999;	99US-0149929P.
PR	23-AUG-1999;	99US-0149902P.
PR	23-AUG-1999;	99US-0149930P.
PR	25-AUG-1999;	99US-0150566P.
PR	26-AUG-1999;	99US-0150884P.
PR	27-AUG-1999;	99US-0151065P.
PR	27-AUG-1999;	99US-0151066P.
PR	27-AUG-1999;	99US-0151080P.
PR	27-AUG-1999;	99US-0151303P.
PR	30-AUG-1999;	99US-0151438P.
PR	31-AUG-1999;	99US-0151930P.
PR	01-SEP-1999;	99US-0152363P.
PR	07-SEP-1999;	99US-0153070P.
PR	10-SEP-1999;	99US-0153758P.
PR	13-SEP-1999;	99US-0154018P.
PR	15-SEP-1999;	99US-0154039P.
PR	16-SEP-1999;	99US-0154779P.
PR	20-SEP-1999;	99US-0155139P.
PR	22-SEP-1999;	99US-0155139P.
PR	06-JUL-1999;	99US-0155486P.
PR	24-SEP-1999;	99US-0155659P.
PR	28-SEP-1999;	99US-0156458P.
PR	29-SEP-1999;	99US-0156596P.
PR	04-OCT-1999;	99US-0157117P.
PR	05-OCT-1999;	99US-0157753P.
PR	06-OCT-1999;	99US-0157865P.
PR	07-OCT-1999;	99US-0158029P.
PR	08-OCT-1999;	99US-0158232P.
PR	12-OCT-1999;	99US-0158369P.
PR	13-OCT-1999;	99US-0159293P.
PR	13-OCT-1999;	99US-0159294P.
PR	13-OCT-1999;	99US-0159295P.
PR	14-OCT-1999;	99US-0159329P.
PR	14-OCT-1999;	99US-0159330P.
PR	14-OCT-1999;	99US-0159331P.
PR	14-OCT-1999;	99US-0159637P.
PR	14-OCT-1999;	99US-0159638P.
PR	18-OCT-1999;	99US-0160741P.
PR	21-OCT-1999;	99US-0160767P.
PR	21-OCT-1999;	99US-0160768P.
PR	21-OCT-1999;	99US-0160770P.
PR	21-OCT-1999;	99US-0160814P.
PR	21-OCT-1999;	99US-0160815P.
PR	22-OCT-1999;	99US-0160980P.
PR	22-OCT-1999;	99US-0160981P.
PR	22-OCT-1999;	99US-0160989P.
PR	25-OCT-1999;	99US-0161404P.
PR	25-OCT-1999;	99US-0161405P.
PR	25-OCT-1999;	99US-0161406P.
PR	26-OCT-1999;	99US-0161359P.
PR	26-OCT-1999;	99US-0161360P.
PR	26-OCT-1999;	99US-0161361P.
PR	28-OCT-1999;	99US-0161920P.
PR	28-OCT-1999;	99US-0161922P.
PR	28-OCT-1999;	99US-0161933P.
PR	29-OCT-1999;	99US-0162142P.
Query Match 30.8%; Score 553; DB 3; Length 367;		
Best Local Similarity 36.3%; Pred. No. 9,7e-53;		
Matches 120; Conservative 65; Mismatches 132; Indels 14; Gaps 4;		
QY	5	LIINCILITGTCGPLLRLTYTGKRIWMFMSFLSTAGFPILIPLLVSPFLSRRGN 64
Db	12	LLIVTYFFLFGSIASSLLAKYFYVCGSSRWSTWQSQAGFLLLLLIYPHYVLKTTT 71
QY	65	RNPNAENKRTKLFMETPLFIASIVIGLTGLDNYLYSYGLAYLVPVSTSSLIIGTOLA 124
Db	72	RRP-----FRTLRH---LIFSVLIGLVGFNFSLFSGTSLYLPVSTSSLLSTQLV 121
QY	125	FNALFALLVKQKTPPFSINAVLLTVGIGILALHSDGPKAKSKKEYVVGFLMTVAA 184
Db	122	FTLILSRIVKQITFSLNLCVLLTSSVLLALDSSKDKFSLGTLTKYFYGYSTICAG 181
QY	185	LIYAFILPLVELTYKKAQOEITFPLVLEIQMVMCIAATFFCVIGMFIYVGDPKVIAREARE 244
Db	182	LLFALYLPVTEKLY---RTVVCYAMMEVQLVMEFAATVFATIGMACGEGFKEMVKEANH 238
QY	245	FKIGGSVFYALIVITGI1-WQGFPLGAIGVFCASSLASGLVLSVLLPVTVEFAVCFER 303
Db	239	VFTKGTFYVTFAILANVVTWLSPAATSGMVLTSGITGTCMTALLAMNVIGGVAYG 298
QY	304	EKFQAEKGVSLLSLWGFVSFYGFYGFSGKK 334
Db	299	DVFGGVKIVSTVLCIWFSSSYTYGMKMKK 329
RESULT 9		
AAG22804		
ID	AAG22804 standard; protein; 382 AA.	
XX		
AC	AAG22804;	

[illegible]

[illegible]



RESULT 11

AC AAG44536;

XX  
DE Arabidopsis thaliana prote

KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.

XX  
PN  
FP1033405-A2

FD 00-321-2000:  
XX  
25 FEB 2000: 2000EB-00301439

PR 25-FEB-1999; 99US-0121823F  
PR 05-MAR-1999; 99US-0123180P.  
26 MAR 1999 00123549D

PR 25-MAR-1999; 99US-0126264F.  
PR 29-MAR-1999; 99US-0126785P.

PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.

PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.

PR	30-APR-1999;	99US-0132407P.
PR	04-MAY-1999;	99US-0132484P.

PR	06-MAY-1999;	99US-0132487P.
PR	07-MAY-1999;	99US-0132863P.

PR 14-MAY-1999; 99US-0134219P  
PR 14-MAY-1999; 99US-0134221P

PR 19-MAY-1999; 99US-0134941P  
PR 20-MAY-1999; 99US-0135124P

PR	21-MAY-1999;	99US-01353531P
PR	22-MAY-1999;	99US-01356293P
PR	24-MAY-1999;	99US-0136021P
PR	25-MAY-1999;	99US-0136382P
PR	27-MAY-1999;	99US-0136782P
PR	28-MAY-1999;	99US-0137222P
PR	01-JUN-1999;	99US-0137528P
PR	03-JUN-1999;	99US-0137502P
PR	04-JUN-1999;	99US-0137724P
PR	07-JUN-1999;	99US-0138094P
PR	08-JUN-1999;	99US-0138540P
PR	10-JUN-1999;	99US-0138847P
PR	10-JUN-1999;	99US-0139119P
PR	14-JUN-1999;	99US-0139452P
PR	16-JUN-1999;	99US-0139453P
PR	17-JUN-1999;	99US-0139492P
PR	17-JUN-1999;	99US-0139454P
PR	18-JUN-1999;	99US-0139455P
PR	18-JUN-1999;	99US-0139457P
PR	18-JUN-1999;	99US-0139458P
PR	18-JUN-1999;	99US-0139460P
PR	18-JUN-1999;	99US-0139461P
PR	18-JUN-1999;	99US-0139462P
PR	18-JUN-1999;	99US-0139463P
PR	18-JUN-1999;	99US-0139750P
PR	18-JUN-1999;	99US-0139763P
PR	21-JUN-1999;	99US-0139817P
PR	21-JUN-1999;	99US-0139893P
PR	23-JUN-1999;	99US-0140353P
PR	23-JUN-1999;	99US-0140354P
PR	24-JUN-1999;	99US-0140695P
PR	24-JUN-1999;	99US-0140823P
PR	28-JUN-1999;	99US-0140991P
PR	28-JUN-1999;	99US-0141287P
PR	30-JUN-1999;	99US-01412842P
PR	01-JUL-1999;	99US-0141254P
PR	01-JUL-1999;	99US-0142055P
PR	02-JUL-1999;	99US-0142390P
PR	06-JUL-1999;	99US-0142803P
PR	08-JUL-1999;	99US-0142920P
PR	09-JUL-1999;	99US-0142977P
PR	12-JUL-1999;	99US-0143342P
PR	13-JUL-1999;	99US-0143362P
PR	14-JUL-1999;	99US-0144181P
PR	15-JUL-1999;	99US-0144005P
PR	16-JUL-1999;	99US-0144085P
PR	16-JUL-1999;	99US-0144086P
PR	19-JUL-1999;	99US-0144331P
PR	19-JUL-1999;	99US-0144332P
PR	19-JUL-1999;	99US-0144333P
PR	19-JUL-1999;	99US-0144334P
PR	19-JUL-1999;	99US-0144335P
PR	20-JUL-1999;	99US-0144352P
PR	20-JUL-1999;	99US-0144352P
PR	20-JUL-1999;	99US-0144684P
PR	21-JUL-1999;	99US-0144814P
PR	21-JUL-1999;	99US-0145086P
PR	21-JUL-1999;	99US-0145088P
PR	22-JUL-1999;	99US-0145085P
PR	22-JUL-1999;	99US-0145087P
PR	22-JUL-1999;	99US-0145089P
PR	22-JUL-1999;	99US-0145192P
PR	22-JUL-1999;	99US-0145214P
PR	23-JUL-1999;	99US-0145218P
PR	23-JUL-1999;	99US-0145224P
PR	23-JUL-1999;	99US-0145276P
PR	26-JUL-1999;	99US-0145913P
PR	27-JUL-1999;	99US-0145918P
PR	27-JUL-1999;	99US-0145919P
PR	27-JUL-1999;	99US-0145951P
PR	28-AUG-1999;	99US-0146386P



us-09-913-767-8.rag

Mon Sep 27 08:51:15 2004

PR	04-MAY-1999;	99US-0132484P.	PR	22-JUL-1999;	99US-0145085P.
PR	05-MAY-1999;	99US-0132485P.	PR	22-JUL-1999;	99US-0145087P.
PR	06-MAY-1999;	99US-0132486P.	PR	22-JUL-1999;	99US-0145089P.
PR	07-MAY-1999;	99US-0132487P.	PR	22-JUL-1999;	99US-0145192P.
PR	08-MAY-1999;	99US-0132863P.	PR	23-JUL-1999;	99US-0145145P.
PR	11-MAY-1999;	99US-0134256P.	PR	23-JUL-1999;	99US-0145218P.
PR	14-MAY-1999;	99US-0134218P.	PR	23-JUL-1999;	99US-0145224P.
PR	14-MAY-1999;	99US-0134219P.	PR	26-JUL-1999;	99US-0145276P.
PR	14-MAY-1999;	99US-0134221P.	PR	27-JUL-1999;	99US-0145913P.
PR	14-MAY-1999;	99US-0134370P.	PR	27-JUL-1999;	99US-0145918P.
PR	18-MAY-1999;	99US-0134768P.	PR	27-JUL-1999;	99US-0145919P.
PR	19-MAY-1999;	99US-0134941P.	PR	28-JUL-1999;	99US-0145951P.
PR	20-MAY-1999;	99US-0135124P.	PR	02-AUG-1999;	99US-0146386P.
PR	21-MAY-1999;	99US-0135353P.	PR	02-AUG-1999;	99US-0146388P.
PR	24-MAY-1999;	99US-0135623P.	PR	02-AUG-1999;	99US-0146389P.
PR	25-MAY-1999;	99US-0136021P.	PR	02-AUG-1999;	99US-0147038P.
PR	27-MAY-1999;	99US-0136392P.	PR	03-AUG-1999;	99US-0147204P.
PR	28-MAY-1999;	99US-0136782P.	PR	04-AUG-1999;	99US-0147204P.
PR	01-JUN-1999;	99US-0137222P.	PR	04-AUG-1999;	99US-0147302P.
PR	03-JUN-1999;	99US-0137528P.	PR	05-AUG-1999;	99US-0147192P.
PR	04-JUN-1999;	99US-0137502P.	PR	05-AUG-1999;	99US-0147260P.
PR	07-JUN-1999;	99US-0137724P.	PR	06-AUG-1999;	99US-0147303P.
PR	08-JUN-1999;	99US-0138094P.	PR	06-AUG-1999;	99US-0147416P.
PR	10-JUN-1999;	99US-0138540P.	PR	09-AUG-1999;	99US-0147493P.
PR	14-JUN-1999;	99US-0139119P.	PR	10-AUG-1999;	99US-0147935P.
PR	16-JUN-1999;	99US-0139452P.	PR	11-AUG-1999;	99US-0148171P.
PR	16-JUN-1999;	99US-0139453P.	PR	11-AUG-1999;	99US-0148319P.
PR	17-JUN-1999;	99US-0139454P.	PR	12-AUG-1999;	99US-0148341P.
PR	18-JUN-1999;	99US-0139455P.	PR	13-AUG-1999;	99US-0148565P.
PR	18-JUN-1999;	99US-0139456P.	PR	13-AUG-1999;	99US-0148684P.
PR	18-JUN-1999;	99US-0139457P.	PR	16-AUG-1999;	99US-0149368P.
PR	18-JUN-1999;	99US-0139458P.	PR	17-AUG-1999;	99US-0149175P.
PR	18-JUN-1999;	99US-0139459P.	PR	18-AUG-1999;	99US-0149426P.
PR	18-JUN-1999;	99US-0139460P.	PR	20-AUG-1999;	99US-0149722P.
PR	18-JUN-1999;	99US-0139461P.	PR	20-AUG-1999;	99US-0149723P.
PR	18-JUN-1999;	99US-0139462P.	PR	20-AUG-1999;	99US-0149829P.
PR	18-JUN-1999;	99US-0139463P.	PR	23-AUG-1999;	99US-0149902P.
PR	18-JUN-1999;	99US-0139750P.	PR	23-AUG-1999;	99US-0149930P.
PR	18-JUN-1999;	99US-0139763P.	PR	25-AUG-1999;	99US-0150566P.
PR	18-JUN-1999;	99US-0139817P.	PR	26-AUG-1999;	99US-0150884P.
PR	21-JUN-1999;	99US-0139899P.	PR	27-AUG-1999;	99US-0151065P.
PR	22-JUN-1999;	99US-0140353P.	PR	27-AUG-1999;	99US-0151066P.
PR	23-JUN-1999;	99US-0140354P.	PR	27-AUG-1999;	99US-0151080P.
PR	24-JUN-1999;	99US-0140695P.	PR	27-AUG-1999;	99US-0151303P.
PR	28-JUN-1999;	99US-0140823P.	PR	30-AUG-1999;	99US-0151438P.
PR	29-JUN-1999;	99US-0140991P.	PR	31-AUG-1999;	99US-0151930P.
PR	30-JUN-1999;	99US-0141287P.	PR	01-SEP-1999;	99US-0151930P.
PR	01-JUL-1999;	99US-0141842P.	PR	07-SEP-1999;	99US-0152363P.
PR	01-JUL-1999;	99US-0142154P.	PR	10-SEP-1999;	99US-0153070P.
PR	02-JUL-1999;	99US-0142055P.	PR	13-SEP-1999;	99US-0153758P.
PR	06-JUL-1999;	99US-0142390P.	PR	15-SEP-1999;	99US-0154018P.
PR	08-JUL-1999;	99US-0142803P.	PR	16-SEP-1999;	99US-0154039P.
PR	09-JUL-1999;	99US-0142920P.	PR	20-SEP-1999;	99US-0154779P.
PR	12-JUL-1999;	99US-0142977P.	PR	22-SEP-1999;	99US-0155139P.
PR	13-JUL-1999;	99US-0143542P.	PR	23-SEP-1999;	99US-0155486P.
PR	14-JUL-1999;	99US-0143624P.	PR	24-SEP-1999;	99US-0155659P.
PR	15-JUL-1999;	99US-0144005P.	PR	28-SEP-1999;	99US-0156458P.
PR	16-JUL-1999;	99US-0144085P.	PR	29-SEP-1999;	99US-0156596P.
PR	16-JUL-1999;	99US-0144325P.	PR	04-OCT-1999;	99US-0157117P.
PR	19-JUL-1999;	99US-0144331P.	PR	05-OCT-1999;	99US-0157533P.
PR	19-JUL-1999;	99US-0144332P.	PR	07-OCT-1999;	99US-0157865P.
PR	19-JUL-1999;	99US-0144333P.	PR	07-OCT-1999;	99US-0158029P.
PR	19-JUL-1999;	99US-0144334P.	PR	08-OCT-1999;	99US-0158232P.
PR	19-JUL-1999;	99US-0144335P.	PR	12-OCT-1999;	99US-0158369P.
PR	20-JUL-1999;	99US-0144336P.	PR	13-OCT-1999;	99US-0159293P.
PR	20-JUL-1999;	99US-0144632P.	PR	13-OCT-1999;	99US-0159294P.
PR	20-JUL-1999;	99US-0144884P.	PR	13-OCT-1999;	99US-0159295P.
PR	21-JUL-1999;	99US-0144814P.	PR	14-OCT-1999;	99US-0159329P.
PR	21-JUL-1999;	99US-0145086P.	PR	14-OCT-1999;	99US-0159330P.
PR	21-JUL-1999;	99US-0145088P.	PR	14-OCT-1999;	99US-0159331P.
PR	21-JUL-1999;	99US-0145088P.	PR	14-OCT-1999;	99US-0159337P.
PR	21-JUL-1999;	99US-0145088P.	PR	14-OCT-1999;	99US-0159638P.
PR	21-JUL-1999;	99US-0145088P.	PR	18-OCT-1999;	99US-0159584P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160741P.
PR	21-JUL-1999;	99US-0145088P.	PR	21-OCT-1999;	99US-0160767P.





Mon Sep 27 08:51:15 2004

PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-014638P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148344P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 23-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 01-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 22-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 28-SEP-1999; 99US-0156596P.  
PR 28-SEP-1999; 99US-0157117P.  
PR 04-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159294P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 18-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.

PR 21-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 25-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 28.6%; Score 513.5; DB 3; Length 386;  
Best Local Similarity 35.1%; Pred. No. 2.7e-48;  
Matches 119; Conservative 71; Mismatches 134; Indels 15; Gaps 6;

Qy 12 ILTIGTCGGPLTLRLYFTNGGKRIFWMSFLSTAGFFIILIPILVSLRRRGRNPNNAE 71  
Db 43 LLLAGETIATLLGRLYEKGKSTWLETVLQVGFPLT-PCYYYL---KPBPSTKTTI 97  
Qy 72 NKRKTKLFMETPLFTASIVIGLTDNLYSYGLAYLPVSTSSLIIGTQLAFNALFAP 131  
Db 98 TKKTTSSFL--TLSLVYIGLLVAGHCILYSGLLYLFVSTFSLISASQLAFNAVFSY 154  
Qy 132 LLVKQKTFPSINAVLLTVGIGILALHSDGDKPAKES---KKEYVVGFLMTVVAALLY 187  
Db 155 FLNSQKITPFLNLSVLLTISTLLVIQHEPSPSTSKAASKYVIGICAVGSSAGY 214  
Qy 188 AFILPLVELTYKKARQETTFPLVLEIQMVMCLAAATFFCVIGMFIIVGDFKVIAREAREFKI 247  
Db 215 SLVSLTDYAFKELKLYTKKATKALDMMATYPSMWATCVVVVGLFGSGGKXKLSTEMEEFQL 274  
Qy 248 GGSVFFYALIVI-TGIWQGFFLGAIIVFCASSLASGLISVLLPTEVPANVCPEKFX 306  
Db 275 GKS--SYILINIGTISWQACLGISVGLIIEVSLFSNVISTLCLPVPVPLAVVFFRDEM 332  
Qy 307 QAEKGVSLLSLWCFVSFYGFGEKSGKKVVDKPPPETE 345  
Db 333 SGIKLVAMFLAINGFVSYGQHYVYNDKRPBEDQELPQSK 371

RESULT 14  
AAG52937 standard; protein; 387 AA.  
XX AAG52937;  
XX 18-OCT-2000 (first entry)  
XX Arabidopsis thaliana protein fragment SEQ ID NO: 67346.  
DE Arabidopsis thaliana.  
KW Protein identification; signal transduction pathway; metabolic pathway;  
KW hybridisation assay; genetic mapping; gene expression control; promoter;  
KW termination sequence.  
XX Arabidopsis thaliana.  
XX EP1033405-A2.  
XX 06-SEP-2000.  
PD 25-FEB-2000; 2000EP-00301439.  
XX 25-FEB-1999; 99US-0121825P.  
PR 05-MAR-1999; 99US-0123180P.  
PR 09-MAR-1999; 99US-0123548P.  
PR 23-MAR-1999; 99US-0125788P.  
PR 25-MAR-1999; 99US-0126264P.  
PR 29-MAR-1999; 99US-0126785P.  
PR 01-APR-1999; 99US-0127462P.

PR 06-APR-1999; 99US-0128234P.  
PR 08-APR-1999; 99US-0128714P.  
PR 16-APR-1999; 99US-0129845P.  
PR 19-APR-1999; 99US-0130077P.  
PR 21-APR-1999; 99US-0130449P.  
PR 23-APR-1999; 99US-0130510P.  
PR 23-APR-1999; 99US-0130891P.  
PR 28-APR-1999; 99US-0131449P.  
PR 30-APR-1999; 99US-0132048P.  
PR 30-APR-1999; 99US-0132407P.  
PR 04-MAY-1999; 99US-0132485P.  
PR 05-MAY-1999; 99US-0132486P.  
PR 06-MAY-1999; 99US-0132486P.  
PR 07-MAY-1999; 99US-0132487P.  
PR 07-MAY-1999; 99US-0132863P.  
PR 11-MAY-1999; 99US-01334256P.  
PR 14-MAY-1999; 99US-01334218P.  
PR 14-MAY-1999; 99US-01334219P.  
PR 14-MAY-1999; 99US-01334221P.  
PR 14-MAY-1999; 99US-0134370P.  
PR 18-MAY-1999; 99US-0134768P.  
PR 19-MAY-1999; 99US-0134941P.  
PR 20-MAY-1999; 99US-0135124P.  
PR 21-MAY-1999; 99US-0135353P.  
PR 24-MAY-1999; 99US-0135629P.  
PR 25-MAY-1999; 99US-0136021P.  
PR 27-MAY-1999; 99US-0136392P.  
PR 28-MAY-1999; 99US-0136782P.  
PR 01-JUN-1999; 99US-0137222P.  
PR 03-JUN-1999; 99US-0137528P.  
PR 04-JUN-1999; 99US-0137502P.  
PR 07-JUN-1999; 99US-0137724P.  
PR 08-JUN-1999; 99US-0138094P.  
PR 10-JUN-1999; 99US-0138540P.  
PR 10-JUN-1999; 99US-0138847P.  
PR 14-JUN-1999; 99US-0139119P.  
PR 16-JUN-1999; 99US-0139452P.  
PR 16-JUN-1999; 99US-0139453P.  
PR 17-JUN-1999; 99US-0139492P.  
PR 18-JUN-1999; 99US-0139454P.  
PR 18-JUN-1999; 99US-0139455P.  
PR 18-JUN-1999; 99US-0139456P.  
PR 18-JUN-1999; 99US-0139457P.  
PR 18-JUN-1999; 99US-0139458P.  
PR 18-JUN-1999; 99US-0139459P.  
PR 18-JUN-1999; 99US-0139460P.  
PR 18-JUN-1999; 99US-0139461P.  
PR 18-JUN-1999; 99US-0139462P.  
PR 18-JUN-1999; 99US-0139463P.  
PR 18-JUN-1999; 99US-0139750P.  
PR 18-JUN-1999; 99US-0139763P.  
PR 21-JUN-1999; 99US-0139817P.  
PR 22-JUN-1999; 99US-0139899P.  
PR 23-JUN-1999; 99US-0140353P.  
PR 23-JUN-1999; 99US-0140354P.  
PR 24-JUN-1999; 99US-0140695P.  
PR 28-JUN-1999; 99US-0140823P.  
PR 29-JUN-1999; 99US-0140991P.  
PR 30-JUN-1999; 99US-0141287P.  
PR 01-JUL-1999; 99US-0141842P.  
PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142820P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143542P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 16-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144325P.  
PR 19-JUL-1999; 99US-0144331P.

PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 19-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 21-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145085P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 28-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 03-AUG-1999; 99US-0146389P.  
PR 04-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 05-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.  
PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.

PR	13-OCT-1999;	99US-0159294P.	PN	EP1033405-A2.
PR	13-OCT-1999;	99US-0159295P.	XX	
PR	13-OCT-1999;	99US-0159329P.	PD	
PR	14-OCT-1999;	99US-0159330P.	XX	06-SEP-2000.
PR	14-OCT-1999;	99US-0159331P.	XX	
PR	14-OCT-1999;	99US-0159637P.	PF	25-FEB-2000; 2000EP-00301439.
PR	14-OCT-1999;	99US-0159638P.	XX	
PR	14-OCT-1999;	99US-0159584P.	PR	25-FEB-1999;
PR	18-OCT-1999;	99US-0160741P.	PR	05-MAR-1999;
PR	21-OCT-1999;	99US-0160767P.	PR	09-MAR-1999;
PR	21-OCT-1999;	99US-0160768P.	PR	23-MAR-1999;
PR	21-OCT-1999;	99US-0160770P.	PR	25-MAR-1999;
PR	21-OCT-1999;	99US-0160814P.	PR	29-MAR-1999;
PR	21-OCT-1999;	99US-0160815P.	PR	01-APR-1999;
PR	21-OCT-1999;	99US-0160980P.	PR	06-APR-1999;
PR	22-OCT-1999;	99US-0160981P.	PR	08-APR-1999;
PR	22-OCT-1999;	99US-0160989P.	PR	16-APR-1999;
PR	22-OCT-1999;	99US-0161404P.	PR	19-APR-1999;
PR	25-OCT-1999;	99US-0161403P.	PR	21-APR-1999;
PR	25-OCT-1999;	99US-0161406P.	PR	23-APR-1999;
PR	26-OCT-1999;	99US-0161359P.	PR	28-APR-1999;
PR	26-OCT-1999;	99US-0161360P.	PR	30-APR-1999;
PR	26-OCT-1999;	99US-0161361P.	PR	30-APR-1999;
PR	28-OCT-1999;	99US-0161920P.	PR	04-MAY-1999;
PR	28-OCT-1999;	99US-0161922P.	PR	05-MAY-1999;
PR	28-OCT-1999;	99US-0161993P.	PR	06-MAY-1999;
PR	29-OCT-1999;	99US-0162142P.	PR	06-MAY-1999;
Query Match 28.6%; Score 513.5; DB 3; Length 387;			PR	07-MAY-1999;
Best Local Similarity 35.1%; Pred.No. 2.7e-48;			PR	11-MAY-1999;
Matches 119; Conservative 71; Mismatches 134; Indels 15; Gaps 6;			PR	14-MAY-1999;
			PR	14-MAY-1999;
QY	12	ILITGCGGPLELTRYFTNGKRIMFMSFLSTAGPPIILIPLLVFLSRRGRNRPNNAE 71	PR	14-MAY-1999;
Db	44	LLLAGETIATLLGRLYEYEGKSTWLETIVQLVGFPLT-PCYYYL---KPEPSKTKTI 98	PR	14-MAY-1999;
QY	72	NKRKYKLFMETPLFIASIVIGLTDNLYSYGLAYLPVSTSLIIGTOLAFNALPAF 131	PR	18-MAY-1999;
Db	99	TKKTTSSFL---TSLVYIGLGLVAGHCILYSFGLLYLPVSTFSLISASQAFNAFVSY 155	PR	20-MAY-1999;
QY	132	LLVKQKFTPFNSAVLLTVGILGALHSDGDKPAKES----KKEYVVGFLMTVVVAALLY 187	PR	21-MAY-1999;
Db	156	FLNSQKITPFLNSVLLTISITLLVIQHEPSPSTSKSAKSKYVIGYICAVGSSAGY 215	PR	24-MAY-1999;
QY	188	AFILPVELTYKARQEIITFPLVLEIQVMVCLAAATFCVIGMFIVGDFKVIAREAREPKI 247	PR	25-MAY-1999;
Db	216	SLVLSLTDVAFEXILKKYTFKAILDMATYPSMVATCVVVVGLFGSGGKKLSTEMEETFOL 275	PR	27-MAY-1999;
QY	248	GGSVFYALIVI-TGIWQGFPLGATGIVFCASSLASGVLISVLLPVTEVPVAVCFREKF 306	PR	28-MAY-1999;
Db	276	GKS--SYILINIGSTISWQACLGISVGLIIEVSSLFNSVISTLCPLVPVPLAVVFRDEM 333	PR	01-JUN-1999;
QY	307	QAEKGVSLLSLWGFVSFYFGEFKSGKVVVDKPPPETE 345	PR	03-JUN-1999;
Db	334	SGIKLVAMFLAIWGFVSFYQHVVDNRKPEDEQLPQSK 372	PR	04-JUN-1999;
RESULT 15			PR	07-JUN-1999;
AAG52936			PR	08-JUN-1999;
ID	AAG52936 standard; protein; 394 AA.		PR	10-JUN-1999;
XX	AAG52936;		PR	14-JUN-1999;
XX	18-OCT-2000 (first entry)		PR	16-JUN-1999;
DT	Arabidopsis thaliana protein fragment SEQ ID NO: 67345.		PR	16-JUN-1999;
XX	Protein identification; signal transduction pathway; metabolic pathway;		PR	18-JUN-1999;
DE	hybridisation assay; genetic mapping; gene expression control; promoter;		PR	18-JUN-1999;
XX	termination sequence.		PR	18-JUN-1999;
KW	Arabidopsis thaliana.		PR	21-JUN-1999;
KW			PR	22-JUN-1999;
OS			PR	23-JUN-1999;
XX			PR	24-JUN-1999;
XX			PR	28-JUN-1999;
XX			PR	29-JUN-1999;
XX			PR	30-JUN-1999;
XX			PR	01-JUL-1999;

PR 01-JUL-1999; 99US-0142154P.  
PR 02-JUL-1999; 99US-0142055P.  
PR 06-JUL-1999; 99US-0142390P.  
PR 08-JUL-1999; 99US-0142803P.  
PR 09-JUL-1999; 99US-0142920P.  
PR 12-JUL-1999; 99US-0142977P.  
PR 13-JUL-1999; 99US-0143422P.  
PR 14-JUL-1999; 99US-0143624P.  
PR 15-JUL-1999; 99US-0144005P.  
PR 16-JUL-1999; 99US-0144085P.  
PR 19-JUL-1999; 99US-0144086P.  
PR 19-JUL-1999; 99US-0144331P.  
PR 19-JUL-1999; 99US-0144332P.  
PR 19-JUL-1999; 99US-0144333P.  
PR 19-JUL-1999; 99US-0144334P.  
PR 20-JUL-1999; 99US-0144335P.  
PR 20-JUL-1999; 99US-0144352P.  
PR 20-JUL-1999; 99US-0144632P.  
PR 20-JUL-1999; 99US-0144884P.  
PR 21-JUL-1999; 99US-0144814P.  
PR 21-JUL-1999; 99US-0145086P.  
PR 22-JUL-1999; 99US-0145088P.  
PR 22-JUL-1999; 99US-0145087P.  
PR 22-JUL-1999; 99US-0145089P.  
PR 22-JUL-1999; 99US-0145192P.  
PR 23-JUL-1999; 99US-0145145P.  
PR 23-JUL-1999; 99US-0145218P.  
PR 23-JUL-1999; 99US-0145224P.  
PR 26-JUL-1999; 99US-0145276P.  
PR 27-JUL-1999; 99US-0145913P.  
PR 27-JUL-1999; 99US-0145918P.  
PR 27-JUL-1999; 99US-0145919P.  
PR 28-JUL-1999; 99US-0145951P.  
PR 02-AUG-1999; 99US-0146386P.  
PR 02-AUG-1999; 99US-0146388P.  
PR 02-AUG-1999; 99US-0146389P.  
PR 03-AUG-1999; 99US-0147038P.  
PR 04-AUG-1999; 99US-0147204P.  
PR 04-AUG-1999; 99US-0147302P.  
PR 05-AUG-1999; 99US-0147192P.  
PR 05-AUG-1999; 99US-0147260P.  
PR 06-AUG-1999; 99US-0147303P.  
PR 06-AUG-1999; 99US-0147416P.  
PR 09-AUG-1999; 99US-0147493P.  
PR 09-AUG-1999; 99US-0147935P.  
PR 10-AUG-1999; 99US-0148171P.  
PR 11-AUG-1999; 99US-0148319P.  
PR 12-AUG-1999; 99US-0148341P.  
PR 13-AUG-1999; 99US-0148565P.  
PR 13-AUG-1999; 99US-0148684P.  
PR 16-AUG-1999; 99US-0149368P.  
PR 17-AUG-1999; 99US-0149175P.  
PR 18-AUG-1999; 99US-0149426P.  
PR 20-AUG-1999; 99US-0149722P.  
PR 20-AUG-1999; 99US-0149723P.  
PR 20-AUG-1999; 99US-0149929P.  
PR 23-AUG-1999; 99US-0149902P.  
PR 23-AUG-1999; 99US-0149930P.  
PR 25-AUG-1999; 99US-0150566P.  
PR 26-AUG-1999; 99US-0150884P.  
PR 27-AUG-1999; 99US-0151065P.  
PR 27-AUG-1999; 99US-0151066P.  
PR 27-AUG-1999; 99US-0151080P.  
PR 30-AUG-1999; 99US-0151303P.  
PR 31-AUG-1999; 99US-0151438P.  
PR 01-SEP-1999; 99US-0151930P.  
PR 07-SEP-1999; 99US-0152363P.  
PR 10-SEP-1999; 99US-0153070P.  
PR 13-SEP-1999; 99US-0153758P.  
PR 15-SEP-1999; 99US-0154018P.  
PR 16-SEP-1999; 99US-0154039P.

PR 20-SEP-1999; 99US-0154779P.  
PR 22-SEP-1999; 99US-0155139P.  
PR 23-SEP-1999; 99US-0155486P.  
PR 24-SEP-1999; 99US-0155659P.  
PR 28-SEP-1999; 99US-0156458P.  
PR 29-SEP-1999; 99US-0156596P.  
PR 04-OCT-1999; 99US-0157117P.  
PR 05-OCT-1999; 99US-0157753P.  
PR 06-OCT-1999; 99US-0157865P.  
PR 07-OCT-1999; 99US-0158029P.  
PR 08-OCT-1999; 99US-0158232P.  
PR 12-OCT-1999; 99US-0158369P.  
PR 13-OCT-1999; 99US-0159293P.  
PR 13-OCT-1999; 99US-0159295P.  
PR 14-OCT-1999; 99US-0159329P.  
PR 14-OCT-1999; 99US-0159330P.  
PR 14-OCT-1999; 99US-0159331P.  
PR 14-OCT-1999; 99US-0159637P.  
PR 14-OCT-1999; 99US-0159638P.  
PR 18-OCT-1999; 99US-0159584P.  
PR 21-OCT-1999; 99US-0160741P.  
PR 21-OCT-1999; 99US-0160767P.  
PR 21-OCT-1999; 99US-0160768P.  
PR 21-OCT-1999; 99US-0160770P.  
PR 21-OCT-1999; 99US-0160814P.  
PR 22-OCT-1999; 99US-0160815P.  
PR 22-OCT-1999; 99US-0160980P.  
PR 22-OCT-1999; 99US-0160981P.  
PR 22-OCT-1999; 99US-0160989P.  
PR 25-OCT-1999; 99US-0161404P.  
PR 25-OCT-1999; 99US-0161405P.  
PR 25-OCT-1999; 99US-0161406P.  
PR 26-OCT-1999; 99US-0161359P.  
PR 26-OCT-1999; 99US-0161360P.  
PR 26-OCT-1999; 99US-0161361P.  
PR 28-OCT-1999; 99US-0161920P.  
PR 28-OCT-1999; 99US-0161922P.  
PR 28-OCT-1999; 99US-0161933P.  
PR 29-OCT-1999; 99US-0162142P.

Query Match 28.6%; Score 513.5; DB 3; Length 394;

Best Local Similarity 35.1%; Pred. No. 2.8e-48;

Matches 119; Conservative 71; Mismatches 134; Indels 15; Gaps 6;

QY 12 ILTIGTCGGPLATRLYFTNGGKRIMPMSFLSTAGFPIILPLLVSLRRGNNPNNAE 71

Db 51 LLAGETIATLGRLYYERGGKSTWLETVLQVLVGFPLT-PCYYL-----KPEFSKTKTI 105

QY 72 NKRKTKFLMETPLFTIASIVIGLLTGLDNLVSYGLAYLPVSTSSLIIGTQLAFNALFAF 131

Db 106 TKKTTSSFL--TSLVYIGLGLLVAGHCILYFGLLYLPVSTFSLISASQLAFNAFVS 162

QY 132 LLVKQKTFPSINAVLLTVGIGILALHSDGDKPAKES-----KKEYVGFGLMTVVAALLY 187

Db 163 FLNSQKITPFIILNSLVLLTISSTLLVIOHEPSPSSTSKSAKSKYVIGVICAAGSSAGY 222

QY 188 AFILPLVELTYKKARQEIFPLVLEIQVMVCLAATFCVIGMFIQVDFKVIAREAEFEKI 247

Db 223 SLVLSLTDYAFEXILKKYTKKAILDMATYPSMVATCVVVGGLFGSGGKKLSTEMEFOQL 282

QY 248 GGSVFYVALIVT-TGIWQGFGLGAGIVFCASSLASGVLISVLLPVTVEFVAVVCREFK 306

Db 283 GKS--SYILINIGSTISWQACILGSLIIEVSSLSFNWISTLCILPVPVLAIVFRDEM 340

QY 307 QAEKGVSLLSLWGFVSFYGFKSGKKVVDKPPPETE 345

Db 341 SGIKLVAMFLAIWGFVSYGQHVWDRKPEEQELPQSK 379

Search completed: September 23, 2004, 23:47:12

Job time : 93 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:47:26 ; Search time 23.6328 Seconds  
(without alignments)  
777.685 Million cell updates/sec

Title: US-09-913-767-8  
Perfect score: 1798  
Sequence: 1 MNGLIINCIIITGTGG.....DKQPPELPILPVSQYVA 356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 389414 seqs, 51625971 residues

Total number of hits satisfying chosen parameters: 389414

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Issued Patents AA.\*  
1: /cgn2\_6/prodata/2/iaa/5A\_COMB.pep.\*  
2: /cgn2\_6/prodata/2/iaa/5B\_COMB.pep.\*  
3: /cgn2\_6/prodata/2/iaa/6A\_COMB.pep.\*  
4: /cgn2\_6/prodata/2/iaa/6B\_COMB.pep.\*  
5: /cgn2\_6/prodata/2/iaa/PCPUS\_COMB.pep.\*  
6: /cgn2\_6/prodata/2/iaa/backfile1.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	122.5	6.8	376	4	US-09-107-532A-6448
2	121	6.7	388	4	US-09-489-039A-9580
3	121	6.7	629	5	PCT-US92-09382-8
4	121	6.7	629	5	US-09-134-000C-4997
5	116	6.5	451	4	US-09-328-352-7659
6	115.5	6.4	419	4	US-09-328-352-4974
7	115	6.4	303	4	US-09-352-991A-27266
8	113.5	6.3	330	4	US-09-543-681A-7116
9	113.5	6.3	342	4	US-09-134-000C-5123
10	113.5	6.3	797	4	US-09-134-000C-4997
11	112	6.2	516	4	US-09-543-681A-5385
12	111.5	6.2	402	4	US-09-107-532A-5693
13	111.5	6.2	463	4	US-09-543-681A-8204
14	111.5	6.2	711	4	US-09-352-991A-32397
15	110.5	6.1	457	4	US-09-107-532A-5465
16	109.5	6.1	506	4	US-09-134-000C-6170
17	109	6.1	434	4	US-09-328-352-7304
18	108	6.0	321	4	US-09-352-991A-31084
19	108	6.0	327	4	US-09-352-991A-25058
20	107.5	6.0	281	4	US-09-540-236-2349
21	107	6.0	457	4	US-09-543-681A-6044
22	107	6.0	713	4	US-09-543-681A-6094
23	106.5	5.9	504	4	US-09-543-681A-6790
24	105.5	5.9	512	4	US-09-107-532A-5262
25	105.5	5.9	514	4	US-09-543-681A-4255
26	105	5.8	311	4	US-09-134-000C-5171
27	105	5.8	826	4	US-09-134-000C-4999

28	104	5.8	329	4	US-09-352-991A-18070	Sequence 18070, A
29	104	5.8	405	4	US-09-489-039A-7574	Sequence 7574, Ap
30	104	5.8	445	4	US-09-352-991A-28655	Sequence 28655, A
31	104	5.8	485	4	US-09-134-000C-6031	Sequence 6031, Ap
32	104	5.8	550	4	US-09-543-681A-6026	Sequence 6026, Ap
33	103.5	5.8	535	2	US-08-677-049-11	Sequence 11, Appl
34	103	5.7	290	4	US-09-134-000C-4893	Sequence 4893, Ap
35	103	5.7	538	4	US-09-107-532A-5563	Sequence 5563, Ap
36	102.5	5.7	349	4	US-09-134-000C-4519	Sequence 4519, Ap
37	102.5	5.7	414	4	US-09-489-039A-12578	Sequence 12578, A
38	102	5.7	444	4	US-09-543-681A-5355	Sequence 5355, Ap
39	102	5.7	448	4	US-09-134-000C-3794	Sequence 3794, Ap
40	102	5.7	628	4	US-09-107-532A-5288	Sequence 5288, Ap
41	101.5	5.6	287	4	US-09-134-000C-5055	Sequence 5055, Ap
42	101.5	5.6	292	2	US-09-024-848-2	Sequence 2, Appli
43	101.5	5.6	292	3	US-09-348-116A-2	Sequence 2, Appli
44	101.5	5.6	296	4	US-09-540-236-2922	Sequence 2922, Ap
45	101.5	5.6	479	4	US-09-489-039A-10571	Sequence 10571, A

ALIGNMENTS

RESULT 1  
US-09-107-532A-6448  
; Sequence 6448, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: LYNN A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneka  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 6448:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 376 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc\_feature  
; LOCATION: (B) LOCATION 1...376  
; SEQUENCE DESCRIPTION: SEQ ID NO: 6448:

US-09-107-532A-6448

Query Match 6.8%; Score 122.5; DB 4; Length 376;  
Best Local Similarity 19.7%; Pred. No. 0.00021;  
Matches 72; Conservative 55; Mismatches 123; Indels 115; Gaps 15;

QY 69 NAENKRTKLFMETPLFIASIVIGLLTGLDNYLSYGLAYL-----PVSTSLI 118  
DB 41 NEKIRRLTVLLFVVFVFLVRSMI--TTILLTFITFYMHLVQIQRVFRIPRTGVLAIM 98  
QY 119 IGTQAFNALFAFLLVKQFTFPFSINAVLLTVGIGILALHSDGD-KPAKESKEYV--- 174  
DB 99 IYTLIVILYLAFA-----TKYIPMLIHQTFMDQSVDFEYHQPDAVDVQLQYIHTYIERN 154  
QY 175 -----VGFLMTV---VAALLYAFILPVELTYKKAR----- 202  
DB 155 DFFEQLONGASIALGYLDQIGMAVAFAMSFILSFFFMIEKKKTIVFSRLFLKSEFSWFF 214  
QY 203 QEI-----TFLVLEIQMWCLAAATFCVIGMFTVGDVKVIAREAREFKIGGSVFY 253  
DB 215 QDIYFADKFVNTFGLVLEAQFVIALNTLLTTIALAVFGHQLLS-----LAIMIFI 267  
QY 254 YALIVITGIWQGFPLGAGIVFCASSLAS-----GVLSIVL----- 290  
DB 268 LSLIPVAGVI-----ISCIPLSFIAYSQGGIRDVVVILITILVHLIBSVYVLPKLMSSK 322  
QY 291 --LPVTEFVAVVCFREKFOAEKGVSLLSLNGFV-----SYFYGEFKSGKVVVDKQPP 342  
DB 323 TELPIFYTFVLLVBERF-----FGVWGLIVGIPITFTFLDLVLK--VKKIPNHLPK 371  
QY 343 ETELP 347  
DB 372 ETDIP 376

RESULT 2  
US-09-489-039A-9580  
; Sequence 9580, Application US/09489039A  
; Patent No. 6610836  
; GENERAL INFORMATION:  
; APPLICANT: Gary Breton et. al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO KLEBSIELLA  
; TITLE OF INVENTION: PNEUMONIAE FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.2004001  
; CURRENT APPLICATION NUMBER: US/09/489,039A  
; CURRENT FILING DATE: 2000-01-27  
; PRIOR APPLICATION NUMBER: US 60/117,747  
; PRIOR FILING DATE: 1999-01-29  
; NUMBER OF SEQ ID NOS: 14342  
; SEQ ID NO 9580  
; LENGTH: 388  
; TYPE: PRT  
; ORGANISM: Klebsiella pneumoniae  
US-09-489-039A-9580

Query Match 6.7%; Score 121; DB 4; Length 388;  
Best Local Similarity 24.9%; Pred. No. 0.00031;  
Matches 71; Conservative 41; Mismatches 93; Indels 80; Gaps 12;

QY 13 LITIGTCGG-PLTLRLYFTNGKRIWPMFSLSTAGFPFIILIPLLVPSLRRRRNRNPNNAE 71  
DB 138 LITANIGGVPAATWLGTTIG-----WRMSFLATAGLLAMVSLWFSLPKGSAGERPD--- 190  
QY 72 NKKRTKLFMETPLFTASIVIGLL--TGLDNYLSYGLAYLPVSTSLIIGTQAFNALFA 130  
DB 191 --VKKELSVLLRPQVLSALLTTVLGAGAMFTLYT-----ISPVLNTLTHASSLP- 238  
QY 131 FLVVKQFTFPFSINAVLLTVGIGILALHSDGDKPAKESKEYVVGFLMTVVAALIYAFI 190  
DB 239 -----ITAMLVL-IGVGFSLGNVLGKFKADRSVSGTLKGLLLMAIML---A 282  
QY 191 LPVLVETKYKARQEIITFPLVLEITQMWCLAAATFCV-----IG 228  
DB 283 IPLLAQSQAGA-----AISMIVWGAATFAVVPPLQMRVVRVAEAPGLSSSVNIG 332

QY 229 MFTVGDPKVIAREAREFKIGGSVFY-----YALIVITGIWQGFFL 269  
DB 333 AFNLGN-----ALGAAAGGAVISGGLGYAFVPMGAIAGLAL 370

RESULT 3  
US-08-132-990A-8  
; Sequence 8, Application US/08132990A  
; Patent No. 5834589  
; GENERAL INFORMATION:  
; APPLICANT: MERUELO, DANIEL  
; APPLICANT: YOSHIMOTO, TAKAYUKI  
; TITLE OF INVENTION: Human Retrovirus Receptor and DNA Coding Therefor  
; NUMBER OF SEQUENCES: 31  
; CORRESPONDENCE ADDRESS: 31  
; ADDRESSEE: Pennie & Edmonds  
; STREET: 1155 Avenue of the Americas  
; CITY: New York  
; STATE: NY  
; COUNTRY: USA  
; ZIP: 10036  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.24  
; CURRENT APPLICATION DATA: US/08/132,990A  
; APPLICATION NUMBER: 08/084,729  
; FILING DATE: 29-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US93/05569  
; FILING DATE: 11-JUN-1993  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/899,075  
; FILING DATE: 11-JUN-1992  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/806,178  
; FILING DATE: 13-DEC-1991  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 07/627,950  
; FILING DATE: 14-DEC-1990  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Mistrock, S. Leslie  
; REGISTRATION NUMBER: 18,872  
; REFERENCE/DOCKET NUMBER: 8105-004-999  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (212) 790-9090  
; TELEFAX: (212) 869-8864  
; TELEX: 66441 PENNIE  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 629 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
US-08-132-990A-8

Query Match 6.7%; Score 121; DB 2; Length 629;  
Best Local Similarity 23.6%; Pred. No. 0.00058;  
Matches 72; Conservative 47; Mismatches 112; Indels 74; Gaps 16;

QY 53 LLVPSLRRRRNRNPNNAENKRTKLFMETPLFTASIVIGLLTGLDNYLSYGLAYLPV 112  
DB 370 LLFKELA-----NVNDRKTKPI-----IATLASGAAVMAFLFD-----LKD 407  
QY 113 STSSLIIGTQAFNALFAFLLVKQFTFPFSINAVLLTVGIGILALHSDGDKPAKE--- 168  
DB 408 LVDLMSIGTLLAYSILVAAACVLV-LRYQEQPNLVQ-----MASTSDLPADQNELA 459



QY 169 SKKEYVVGFLMTVAALLYAFILPLVELTYKKARQETTFPLVLEI--QMVNCLAAATFFCV 226  
Db 460 STNDSQLGFLPEAMFSLKTLSP-----KNMEPSKISGLIWNISTSLIAVLIIT-FCI 512  
QY 227 IGMFIVGDFKVIAREA-----REFKIGGSVFYALIVITIGIWOQFFLGAIGIVFCAS 279  
Db 513 V-----TVLGREALTGKALWAVFLLAGSALLCA--VVTGVIRQ-----PESKT 554  
QY 280 SLASGVLSVLLPVTVEFAVVCFFKFAKGVSLLSLW---GFVSYP-YGFEKSGKKV 335  
Db 555 KLSFKVPFLPVLPILSIFVNVYL--MMQLDQGTWVRFAVWMLIGFIIFYGYGLWHSEAS 612  
QY 336 VDKPQ 340  
Db 613 LDADQ 617

## RESULT 4

PCT-US92-09382-8  
; Sequence 8, Application PC/TUS9209382  
; GENERAL INFORMATION:  
; APPLICANT: MERUELO, DANIEL  
; APPLICANT: YOSHIMOTO, TAKAYUKI  
; TITLE OF INVENTION: Human Retrovirus Receptor and DNA Coding  
; TITLE OF INVENTION: Therefor  
; NUMBER OF SEQUENCES: 8  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: Browdy and Neimark  
; STREET: 419 Seventh Street, N.W.  
; CITY: Washington  
; STATE: DC  
; COUNTRY: USA  
; ZIP: 20004  
; COMPUTER READABLE FORM:  
; MEDIUM TYPE: Floppy disk  
; COMPUTER: IBM PC compatible  
; OPERATING SYSTEM: PC-DOS/MS-DOS  
; SOFTWARE: Patent In Release #1.24  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: PCT/US92/09382  
; FILING DATE: 19921213  
; CLASSIFICATION:  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Livnat, Shmuel  
; REGISTRATION NUMBER: 33,949  
; REFERENCE/DOCKET NUMBER: MERUELO=1  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: 202 628-5197  
; INFORMATION FOR SEQ ID NO: 8:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 629 amino acids  
; TYPE: AMINO ACID  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
PCT-US92-09382-8

Query Match 6.7%; Score 121; DB 5; Length 629;  
Best Local Similarity 23.6%; Pred. No. 0.00058;  
Matches 72; Conservative 47; Mismatches 112; Indels 74; Gaps 16;

QY 53 LLVFSLSRRGNPNNAENKRTKLFLMETPLFIASIVIGLITGLDNYLSYGLAYLPV 112  
Db 370 LLFKFLA-----NVDRTKTPI-----IATLASGAAVAVMLFD-----LKD 407  
QY 113 STSSLLIGTQLAFNALFAFLVVKQKTFPFSINAVLLTVGIGILALHSDGDKPAK----- 168  
Db 408 LVDLMSIGTLAYSLAACVLLV-LRYQEQPNLVYQ-----MASTSDLPADQNELA 459  
QY 169 SKKEYVVGFLMTVAALLYAFILPLVELTYKKARQETTFPLVLEI--QMVNCLAAATFFCV 226  
Db 460 STNDSQLGFLPEAMFSLKTLSP-----KNMEPSKISGLIWNISTSLIAVLIIT-FCI 512  
QY 227 IGMFIVGDFKVIAREA-----REFKIGGSVFYALIVITIGIWOQFFLGAIGIVFCAS, 279

Db 513 V-----TVLGREALTGKALWAVFLLAGSALLCA--VVTGVIRQ-----PESKT 554  
QY 280 SLASGVLSVLLPVTVEFAVVCFFKFAKGVSLLSLW---GFVSYP-YGFEKSGKKV 335  
Db 555 KLSFKVPFLPVLPILSIFVNVYL--MMQLDQGTWVRFAVWMLIGFIIFYGYGLWHSEAS 612  
QY 336 VDKPQ 340  
Db 613 LDADQ 617

## RESULT 5

US-09-328-352-7659  
; Sequence 7659, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 7659  
; LENGTH: 451  
; TYPE: PRT  
; ORGANISM: Acinetobacter baumannii  
US-09-328-352-7659

Query Match 6.5%; Score 116; DB 4; Length 451;  
Best Local Similarity 19.3%; Pred. No. 0.0012;  
Matches 65; Conservative 54; Mismatches 127; Indels 90; Gaps 13;  
QY 17 TCGGPLLTLRYPTNGGKRWFMSFLSTAGPPIILPLLV-----SFLSRRGRNPN 68  
Db 163 SCGGVLSAYLESYG----WAKVIIIGSIPILVIALLLKLPSTQYLITRHQDKAQR 218  
QY 69 NAEN-----KRKTKLFM-----ETPLFIASIVIGLITGLDNYLSYGLAYLPVSTS 115  
Db 219 ILENIQGYSFQKPIKLVLQTAETASDENPV--KVVLG-----KYLWSSMLWLCFTS 269  
QY 116 SLIIGTQLAFNALFAFLVVKQKTFPFSINAVLL-----TVGIGILALHSDGDKPAK 167  
Db 270 LLV-----FYLLTSWPTILKTAGFTSQFSLIAAIFPPGGVIGATIMGWMDKLNPTT 323  
QY 168 ESKEYVVGFLMTVAALLYAFILPLVELTYKKARQETTFPLVLEIOMVWMLAATPCVI 227  
Db 324 VIKYSYLIAPVLFIIAGLVSS-----NIFLLGLTIF-LI 356  
QY 228 GMPVGDVKVIAREAREFKIGGSVFYALIVITIGIWOQ-----GFFLGAI--GIVFCAS 279  
Db 357 GALLAGAQSLLPLA-----AMFYPAVCRVGVSWMIGIGRIGAILGAFFGSLIFTN 409  
QY 280 SLASGVLSVLLPVTVEFAVVCFFKFAKGVSLLS 315  
Db 410 LSLSGIFFILAIPTFISFIALSLKVIYKSKHKQVL 445

## RESULT 6

US-09-328-352-4974  
; Sequence 4974, Application US/09328352  
; Patent No. 6562958  
; GENERAL INFORMATION:  
; APPLICANT: Gary L. Breton et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO ACINETOBACTER  
; TITLE OF INVENTION: BAUMANNII FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: GTC99-03PA  
; CURRENT APPLICATION NUMBER: US/09/328,352  
; CURRENT FILING DATE: 1999-06-04  
; NUMBER OF SEQ ID NOS: 8252  
; SEQ ID NO 4974  
; LENGTH: 419

167 KRSGATDMLRLICWVSLVPPPLPILALSYLFEGP-----QAIERAVLNMSWGG--FCALL 218  
257 IVITGIWOGFFLGAIGIVFCASSLASGVLSVLLPVTEVFA 298  
219 YIAFGATTGFAIWSFLLRHRYPASLVTFFALAV--PVSQMLA 258

RESULT 8  
US-09-543-681A-7116  
; Sequence 7116, Application US/09543681A  
; Patent No. 6605709  
; GENERAL INFORMATION:  
; APPLICANT: GARY BRETON  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS  
; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 2709.1002-001  
; CURRENT APPLICATION NUMBER: US/09/543,681A  
; CURRENT FILING DATE: 2000-04-05  
; PRIOR APPLICATION NUMBER: US 60/128,706  
; PRIOR FILING DATE: 1999-04-09  
; NUMBER OF SEQ ID NOS: 8344  
; SEQ ID NO 7116  
; LENGTH: 330  
; TYPE: PRT  
; ORGANISM: Proteus mirabilis  
US-09-543-681A-7116

Query Match 6.3%; Score 113.5; DB 4; Length 330;  
Best Local Similarity 21.5%; Pred. No. 0.0015;  
Matches 69; Conservative 52; Mismatches 147; Indels 53; Gaps 11;

Qy 2 KNGLIILNCTITGTCGPLLRLTYFTNGGKRIWFMSPFLSTAGFPILIPLLVSLR 61  
Db 31 QSGVLLAICSYILWGIT--PLFYRL--LPGAQPLELLAQRLIWSIPLLLVRLF 80  
Qy 62 RGNRNPNNAENKRTKFLMETPLFTIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIGT 121  
Db 81 -----INKRTWRAVLQDRRSIFMCLFSSLAMAVSWCTFTYALTHQVLAASLYFI 132  
Qy 122 QLAFNALFALLVKKQKTPFSINAVLLTVGIGLALHSDGDKPAKESKKEYVVGFLMTV 181  
Db 133 NPLFSIVLGVIFLKEKLNPAEKNAVILICAGVG-YQLWMTGQLP----- 175  
Qy 182 VAALLYAFILPLVELTYKKARQETTFPLVLEIQVMCLAAATFFCVIGMFIKVIARE 241  
Db 176 ILALVMGAFAPAIYGLIRKFRFVITSFTIETLWLPVA-----IGAMI----WLIMRD 225  
Qy 242 AREFKIGGSV--FYVALIVITGIIWQGFILGAIGIVFCASSLASGVLSVLLPVTE-VFA 298  
Db 226 ESAPSSADNLTRFYIILTAPVTILPLLLFTTAAIK-----RTTLTVIGLAQYIEPTMQFLA 281  
Qy 299 VVCFREKFOAEKGVSLLSLW 319  
Db 282 VFLFHEAPDAVGVSFSL-IW 301

RESULT 9  
US-09-134-000C-5123  
; Sequence 5123, Application US/09134000C  
; Patent No. 6617156  
; GENERAL INFORMATION:  
; APPLICANT: Lynn Doucette-Stamm et al  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 032796-032  
; CURRENT APPLICATION NUMBER: US/09/134,000C  
; CURRENT FILING DATE: 1998-08-13  
; PRIOR APPLICATION NUMBER: US 60/055,778  
; PRIOR FILING DATE: 1997-08-15  
; NUMBER OF SEQ ID NOS: 6812  
; SOFTWARE: PatentIn version 3.1  
; SEQ ID NO 5123  
; LENGTH: 342

13 LTIGTCGG-PILTRLYFTNGGKRI-WFMSFLSTAGFPILIPLLVSLRNRNPN-N 69  
165 LTIANIGVPLATWV-----QNCIGWRMSFLAISVIGLITMLAKWALPQGMVAQKENVK 219  
70 AENKRTKFLMETPLFTIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIGTQLAFNALP 129  
220 AELK-----VLTRTEVLLALLTVLGAAMFTLY-----IAPS-----LT 256  
Qy 130 AFLVKKQKTPFSINAVLLTVGIGLALHSDGDKPAKESKKEYVVGFLMTVVAALLYAF 189  
Db 257 EPHASPTFITF-----MLVLIGVGSIGNHLGGRFADLSINKTLIGLPLVLLIWM----- 307  
Qy 190 ILPLVELTYKKARQBITPPLVLE-----IQVMCLAAATFFCV--IGMFIKVIARE 241  
Db 308 -----VTFPILAQSQIGAAIALVINGAATFALVPLQMRVMS-----VAHE 348  
Qy 242 A-----REFKIGGSVFFYVALIVITGIIWQGFILGAIGIVFCASSLASGV-LISVLL 291  
Db 349 APGLASSVNIAGFNLGNVGAAGALVLDLW-----GYSVSPAGALLAGLLLVLF 402  
Qy 292 PV 293  
Db 403 QI 404

RESULT 7  
US-09-252-991A-27266  
; Sequence 27266, Application US/09252991A  
; Patent No. 6551795  
; GENERAL INFORMATION:  
; APPLICANT: Marc J. Rubenfield et al.  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS  
; TITLE OF INVENTION: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS  
; FILE REFERENCE: 107196.136  
; CURRENT APPLICATION NUMBER: US/09/252,991A  
; CURRENT FILING DATE: 1999-02-18  
; PRIOR APPLICATION NUMBER: US 60/074,788  
; PRIOR FILING DATE: 1998-02-18  
; PRIOR APPLICATION NUMBER: US 60/094,190  
; PRIOR FILING DATE: 1998-07-27  
; NUMBER OF SEQ ID NOS: 33142  
; SEQ ID NO 27266  
; LENGTH: 303  
; TYPE: PRT  
; ORGANISM: Pseudomonas aeruginosa  
US-09-252-991A-27266

Query Match 6.4%; Score 115; DB 4; Length 303;  
Best Local Similarity 23.0%; Pred. No. 0.00094;  
Matches 65; Conservative 52; Mismatches 103; Indels 62; Gaps 14;

Qy 35 IIFWMSF-----LSTAGFPIL-----IPLLVSLRNRNPNNAENKRTKFLP 80  
Db 21 IWGLNFVVIKVGDFPPLFCALFALAAFLVL-----RGP----- 59  
Qy 81 METPLFTIASIVIGLTLGDLNLYSYGLAY-LPVSTSSLIIGTQLAFNALFALLVKKQF 138  
Db 60 MPAP-FGRILLIGVLLGVKFGLLFVGMAGMPAGLSVLOSQVFTTILIAALWIGERP 118  
Qy 139 TPFISINAVLLTVGIGLALHSDGDKPAKESKKEYVVGFLMTVVAALLYAFILPLVELTY 198  
Db 119 SPRLFGLSLAAGLVLLIGL-----ERPLGDS-----LIALVVAALAWSP-----ANIAT 166  
Qy 199 KKARQBITPPLVLEIQMV--MCLAAATFFCVIGMFIKVIAREAREFKIGGSVFFYAL 256

```

; TYPE: PR1
; ORGANISM: Enterococcus faecalis
US-09-134-000C-5123

Query Match
Best Local Similarity 23.8%; Score 113.5; DB 4; Length 342;
Matches 82; Conservative 52; Mismatches 99; Indels 111; Gaps 24;

Qy 11 ILITIG-----TCGPLLTRLRYFTNGGKRIFWMSPLSTAGPPIILIPLL----- 54
   : : ||| : : : : : : : : : : : : : : : : : : : : : : : :
Db 51 VVATIGPVDGNEVWLLTAGGAFSPPY-----WYASLFS--GYVLLFTLIFGLII 101
   : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 55 --VSPLRRRGRRNPNAENKRTKFLMETPLFIASIVIGLLTGLDNLVSYGLAIVP 112
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 102 RGVSFPERH-----NNPEGKRRR--MWNWTLSIGSFVLP-----EFFGILFI-- 141
   : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 113 STSLIIIGTQLAFN-----ALFAPFLVKKQKTFPFSINAVLLTV-----GIGILAHSDGD 163
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 142 ---SLVQGMPLDANGNMHAQFT-----DYFNLFSGVGVALTLLCYLHGNNVIALKTEG- 192
   : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 164 KPAKESKEY---VVGFL---MTVVAALLYAFILPLVELTYKKARQEIPTPLVLEIQWM 217
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 193 -PIRERARNYABILYGLVYTLGVVFAVLMY-----FKTDYFENKFAVTL-----IL 237
   : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 218 CLAAATPFCVGMFTVGDVPKVIAREAREFKIGSVFYVALIVITGIIWGGPF-----LGAIG 273
   : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 238 TLAIVLTVIAN--VGVPK--RKEMLAFLASG-----LTLVVLVALLPSGLEFPRVMIGSE 289
   : : : : : : : : : : : : : : : : : : : : : : : : : :

Qy 274 ---IVFCASS-----LASGVLSVL---LPVTEVFAVVCREK 305
   : : : : : : : : : : : : : : : : : : : : : : : : : :

Db 290 FDLLIKDATSPYTKLMTWISILSPFLVPYPT--AWSYIIPKR 332
   : : : : : : : : : : : : : : : : : : : : : : : : : :

```

RESULT 10  
 US-09-134-000C-4997  
 ; Sequence 4997, Application US/09134000C  
 ; Patent No. 6617156  
 ; GENERAL INFORMATION:  
 ; APPLICANT: Lynn Doucette-Stamm et al  
 ; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
 ; TITLE OF INVENTION: ENTEROCOCCUS FAECALIS FOR DIAGNOSTICS AND THERAPEUTICS  
 ; FILE REFERENCE: 032796-032  
 ; CURRENT APPLICATION NUMBER: US/09/134.000C  
 ; CURRENT FILING DATE: 1998-08-13  
 ; PRIOR APPLICATION NUMBER: US 60/055,778  
 ; PRIOR FILING DATE: 1997-08-15  
 ; NUMBER OF SEQ ID NOS: 6812  
 ; SOFTWARE: PatentIn version 3.1  
 ; SEQ ID NO 4997  
 ; LENGTH: 797  
 ; TYPE: PRT  
 ; ORGANISM: Enterococcus faecalis  
 US-09-134-000C-4997  
 Query Match 6.3%; Score 113.5; DB 4; Length 797;  
 Best Local Similarity 20.9%; Pred. No. 0.0047;  
 Matches 77; Conservative 54; Mismatches 121; Indels 117; Gaps 17;

Qy	23	LTRLFYFTNGGKRIWF	-----MSFLSTAGPPI-----	ILIPLL	54
		:	:	:	:
Db	39	LTLTKIGSAGLSFWFLAKQTFKPKWGHVTLVSVALMSFIVAHSELIMLWDAFYILPLV	98		
Qy	55	VSLSRERGRNRPNNAENKRKTKLFMETPLFIASIVIGLTLGLDNVLYSYG	-----	106	
Db	99	ILGHRIMLDQKRP	-----TLLEFVSVELFITNYFGEFMIGLFSLEYFYFARTFTDQWR	150	
Qy	107	-----LAYLPVSTSSLIIGTQAFNALFAFLVLK-----	OKFTPFSSINAVVLLTV-----	151	
		:	:	:	:
Db	151	YKSRIYAF	-----TTSLLAGGASMIWLPVAVLDLRTNGETLSEITFTFKTEAFALDIIMKN	207	
Qy	152	GIGILAHUSDGDKPAKESKKEYVVGFLMTVVVAALLYAFILPLVBELTYKKARQETFPFL	--	209	
		:	:	:	:
Db	208	MIGVYDTTKYGSIP	-----FIYIGLPLFLCFYFVTKV--	PLKN	246
		:	:	:	:

Qy	210	-VLEIQVMCLAAATPCV-IGMFVGDPKVIAREAREPKIGSVFYALIVITGIWQGF	267
Db	247	KLFGSLFVLLSASFYITPLNLPWHG---MEAPNMFLFRY-SFLSFLVILLAGYGDWKF	302
Qy	268	FLGAIGTVFCASSLASGLVISVLLPVTBVFVAVFCPRKFOAEKG-----VSLLLSL	318
Db	303	EKDDLGLV-----SGLILILL---AIFALA-----MGTKGATSYTYVTLTSPVLTA	345
Qy	319	WGFVSVEFG	327
Db	346	TFLLLYFFG	354

  

RESULT 11	
US-09-543-681A-5385	
Sequence 5385, Application US/09543681A	
Patent No. 6605709	
GENERAL INFORMATION:	
APPLICANT: GARY BRETON	
TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEIN	
TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS	
FILE REFERENCE: 2709.1002-001	
CURRENT APPLICATION NUMBER: US/09/543,681A	
CURRENT FILING DATE: 2000-04-05	
PRIOR APPLICATION NUMBER: US 60/128,706	
PRIOR FILING DATE: 1999-04-09	
NUMBER OF SEQ ID NOS: 8344	
SEQ ID NO 5385	
LENGTH: 516	
TYPE: PRT	
ORGANISM: Proteus mirabilis	
US-09-543-681A-5385	

Query Match	6.2%;	Score 112;	DB 4;	Length 516;	
Best Local Similarity	20.3%;	Pred. No. 0.0038;			
Matches	61;	Conservative 47;	Mismatches 110;	Indels 82;	Gaps 11
Qy	50	LIPLLVSLRRRRNPNNAENKRTKLFLETPFLFIASIVIGLLTGL---DNYLYSYGL	107		
Db	129	LIPLSQSLLN-----NYPKRNMALALWSVTIVVAPILGPLTGGYISDNYHWGW-I	180		
Qy	108	AYLPVSTSSLLIGTLQALNALFALLVKQKTPFSINAVLLTVCIGILALHSDGDKPAK	167		
Db	181	FFINVPFGLII-----MCISNTLAGRETKEIKPIDTIGLVLLVVGVALQIMLDQCKELD	237		
Qy	168	ESKEYVVGFLMTVVAALLYAFIL-----PLVELTYKARQEIITPFLVEIQWVACL	219		
Db	238	WPNSTELL--VLTIIVAVLSFLIWELTDSEHPVIDLSLFSRN-----FTIGCL	285		
Qy	220	AATFCFCVGMFIVGDFKVIAREBEFKIGSVFYVALIVITGIHWQGFPGAIGVFCAS	279		
Db	286	TLSLAYMI-----YFGTIVLPLLLQEVF---GYTATWA	316		
Qy	280	SLAS---GVLISVLLPYTE-----VPAVVCFRKPFQAEKGVSLLLSLW	319		
Db	317	GLAAASVGLLPLIITTIIGKFGGKVDIRIYISFSPIMEAVGCFYMRAYTFEPGMDPATVAM	376		

RESULT 12  
US-09-107-532A-5693  
; Sequence 5693, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ; ; ; ; ;  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSER: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts

```

; COUNTRY: USA
; ZIP: 02354
; COMPUTER READABLE FORM:
; MEDIUM TYPE: CD/ROM ISO9660
; COMPUTER: PC
; OPERATING SYSTEM: <Unknown>
; SOFTWARE: ASCII
; CURRENT APPLICATION DATA:
; APPLICATION NUMBER: US/09/107,532A
; FILING DATE: 30-Jun-1998
; PRIOR APPLICATION DATA:
; APPLICATION NUMBER: 60/085,598
; FILING DATE: 14 May 1998
; APPLICATION NUMBER: 60/051571
; FILING DATE: July 2, 1997
; ATTORNEY/AGENT INFORMATION:
; NAME: Ariniello, Pamela Deneske
; REGISTRATION NUMBER: 40,489
; REFERENCE/DOCKET NUMBER: GTC-012
; TELECOMMUNICATION INFORMATION:
; TELEPHONE: (781)893-5007
; TELEFAX: (781)893-8277
; INFORMATION FOR SEQ ID NO: 5693:
; SEQUENCE CHARACTERISTICS:
; LENGTH: 402 amino acids
; TYPE: amino acid
; TOPOLOGY: linear
; MOLECULE TYPE: protein
; HYPOTHETICAL: YES
; ORIGINAL SOURCE:
; ORGANISM: Enterococcus faecium
; FEATURE:
; NAME/KEY: misc feature
; LOCATION: (B) LOCATION 1...402
; SEQUENCE DESCRIPTION: SEQ ID NO: 5693:
US-09-107-532A-5693

Query Match 6.2%; Score 111.5; DB 4; Length 402;
Best Local Similarity 20.3%; Pred. No. 0.0031;
Matches 64; Conservative 55; Mismatches 101; Indels 95; Gaps 16;

QY 6 IINCIL--TGTGCGPLTLRLY-----FTNGGKRIRWMSFLSTAGF 46
DB 49 LVGLFVLLRTTGTAVSVLAKMFDHLLAFAGQGVDFVAGVANKGSAPFFLSVL---M 104
QY 47 PILIPLVLSFLSRRGRNPNNAENKRXKTLFLMETPLFIASIVIGL--LTGLDNYLYS 104
DB 105 PIVFISAIGILR-----YKILPLFMKAVGLGLSKINGMGK-LES 144
QY 105 YGLAYLPVSTSSLIIGTQAFNAL----PAILLVKQKFTPFSINAVLLTVGIGILALHSD 161
DB 145 YN-----GVASAILQGSVFISIKKELPFLSEKRLFT-MSVSA--MSTVMSIV----- 190
QY 162 GDKPAKESKKEYVVVGLMTVVAALLYAFIL-PLV-----ELTYKKARQEIITFPLVLEIQ 214
DB 191 GSNMALIDSKYVITAVLNLFGYIILASIVNPVLEEKEDELIEENKEQ----- 240
QY 215 MVNCLAAATPCVIGMFIIVGDFKVIAREAREFKIGSVFYVALIVITGIWQGF-----FLG 270
DB 241 -----TFPQMLGEVLDFGHVAITVAAML-----IGFVALIAMINAIHFHIGITFOE 288
QY 271 AIGIVFCASSLASGV 285
DB 289 ILGYIFAPLAFISGI 303

RESULT 13
US-09-543-681A-8204
; Sequence 8204, Application US/09543681A
; Patent No. 6605709
; GENERAL INFORMATION:
; APPLICANT: GARY BRETON
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PROTEUS MIRABILIS

; TITLE OF INVENTION: DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 2709,1002-001
; CURRENT APPLICATION NUMBER: US/09/543,681A
; CURRENT FILING DATE: 2000-04-05
; PRIOR APPLICATION NUMBER: US 60/128,706
; PRIOR FILING DATE: 1999-04-09
; NUMBER OF SEQ ID NOS: 8344
; SEQ ID NO 8204
; LENGTH: 463
; TYPE: PRT
; ORGANISM: Proteus mirabilis
US-09-543-681A-8204

Query Match 6.2%; Score 111.5; DB 4; Length 463;
Best Local Similarity 26.0%; Pred. No. 0.0037;
Matches 84; Conservative 45; Mismatches 99; Indels 95; Gaps 26;

QY 28 FTNGGKRIRWMSFLSTAGPPI--ILPL--LVSFLSRRGRNPNNAENKRXKTKFLMET 83
DB 45 FQSGNQ--WFSTSL--GFLVTGVLPLFLTLTVAIRGRGER-----LSIDL 86
QY 84 PLFIAS---IVIGLLTG-----LDNYLSYGLAYLPV-----STSSLIIGTQAFNA 127
DB 87 PSWFAVLFWIALYLIIVGSTFAMPRVINTAYEMG--FLPLGLIEKNTTHTLTF--LIFNL 142
QY 128 LFAFLLVKQ-----KF--TPFSINAVLLTVGIGILALHSDGDKPAKESKEYVV--GF 177
DB 143 LSNFFMLKQGTWISAIGKFWMTALL--ILLVVGVIAVAKPL--SPIERTPTGLYAVNGF 197
QY 178 L-----MTVVALLYAFILPLVELTYKKARQEIITFP-----LVLEIQM--VMCLAA 221
DB 198 FSGIIDGYQTMVDLSAMAFGGI--VARALYTKG---ITNPKQIGFITIKAGMISVLLAA 252
QY 222 TFFCVIGMFIVG---DFKVIAREAREFKIGSVFYVALIVITGII--WQGFIFGAIIV 275
DB 253 LYLC---LFYLGATSHAVSVFADPALNATNGGQIFSRVVDALFGSVGTW---LMG--GIV 304
QY 276 FCASSLASGVLSVLLPTEVFA 298
DB 305 LLASMT---TLVGVTSAAADYFA 324

RESULT 14
US-09-252-991A-32397
; Sequence 32397, Application US/09252991A
; Patent No. 6551795
; GENERAL INFORMATION:
; APPLICANT: Marc J. Rubenfield et al.
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO PSEUDOMONAS
; FILE REFERENCE: AERUGINOSA FOR DIAGNOSTICS AND THERAPEUTICS
; FILE REFERENCE: 107196.136
; CURRENT APPLICATION NUMBER: US/09/252,991A
; CURRENT FILING DATE: 1999-02-18
; PRIOR APPLICATION NUMBER: US 60/074,788
; PRIOR FILING DATE: 1998-02-18
; PRIOR APPLICATION NUMBER: US 60/094,190
; PRIOR FILING DATE: 1998-07-27
; NUMBER OF SEQ ID NOS: 33142
; SEQ ID NO 32397
; LENGTH: 711
; TYPE: PRT
; ORGANISM: Pseudomonas aeruginosa
US-09-252-991A-32397

Query Match 6.2%; Score 111.5; DB 4; Length 711;
Best Local Similarity 22.1%; Pred. No. 0.0065;
Matches 74; Conservative 42; Mismatches 98; Indels 121; Gaps 16;

QY 14 TIGTC-GGPELLTRLYFTNGGKRIRWMSFLSTAGPPIILPIL-----VSFLSER----- 61
DB 424 TLGSALGGLLAARMVVPALG---WESVLLGGGLPLASLPLLLWACLPSVRFIARAPDA 479
QY 62 -----RGRNPNPN-----AENKRXKTKL-----FLMETPLFIASIVI 92
```

Db 480 ARRLRLGRUPDGWAGSLLEPADESDGTSLLRQILGAELRGGTLLWATFFMGLLII 539  
QY 93 GLLTGLDNYLSYGLAYLPVSTSSLIIGTQAFNALPAFLVKQKFTPFPSINAVLLTVG 152  
Db 540 YLLTN-----WLP-----TLIGTGFSLG-----EATISAMFQLGGT 572  
QY 153 IGHIALHSDGDK--PAKESKEYVVGFLMTV--VAALLYAFILPLVELTYKKARQEIFPL 209  
Db 573 LGALLGSAMDRFDARHVLSLAYVGGALFTLGIASLVHSPAL----- 614  
QY 210 VLRIQVMCLAAATFFCVIGMFIVGDFKVIAREAREFKIGSV-----PYALIVITGIWQ 265  
Db 615 -----LALCVAGVGFCISGS-----QVGANALAADFYPTSRATGVSW- 652  
QY 266 GFFLGAIGIVFCASSLASGLVLSVLLPVTBVFVAV 300  
Db 653 ALGLGRIGSI--VGSLSGALLGLGLGFSGLALL 685

RESULT 15  
US-09-107-532A-5465  
; Sequence 5465, Application US/09107532A  
; Patent No. 6583275  
; GENERAL INFORMATION:  
; APPLICANT: Lynn A Doucette-Stamm and David Bush  
; TITLE OF INVENTION: NUCLEIC ACID AND AMINO ACID SEQUENCES RELATING TO  
; ENTEROCOCCUS FAECIUM FOR DIAGNOSTICS AND THERAPEUTICS  
; NUMBER OF SEQUENCES: 7310  
; CORRESPONDENCE ADDRESS:  
; ADDRESSEE: GENOME THERAPEUTICS CORPORATION  
; STREET: 100 Beaver Street  
; CITY: Waltham  
; STATE: Massachusetts  
; COUNTRY: USA  
; ZIP: 02354  
; COMPUTER-READABLE FORM:  
; MEDIUM TYPE: CD-ROM ISO9660  
; COMPUTER: PC  
; OPERATING SYSTEM: <Unknown>  
; SOFTWARE: ASCII  
; CURRENT APPLICATION DATA:  
; APPLICATION NUMBER: US/09/107,532A  
; FILING DATE: 30-Jun-1998  
; PRIOR APPLICATION DATA:  
; APPLICATION NUMBER: 60/085,598  
; FILING DATE: 14 May 1998  
; APPLICATION NUMBER: 60/051571  
; FILING DATE: July 2, 1997  
; ATTORNEY/AGENT INFORMATION:  
; NAME: Ariniello, Pamela Deneke  
; REGISTRATION NUMBER: 40,489  
; REFERENCE/DOCKET NUMBER: GTC-012  
; TELECOMMUNICATION INFORMATION:  
; TELEPHONE: (781)893-5007  
; TELEFAX: (781)893-8277  
; INFORMATION FOR SEQ ID NO: 5465:  
; SEQUENCE CHARACTERISTICS:  
; LENGTH: 457 amino acids  
; TYPE: amino acid  
; TOPOLOGY: linear  
; MOLECULE TYPE: protein  
; HYPOTHETICAL: YES  
; ORIGINAL SOURCE:  
; ORGANISM: Enterococcus faecium  
; FEATURE:  
; NAME/KEY: misc feature  
; LOCATION: (B) LOCATION 1...457  
; SEQUENCE DESCRIPTION: SEQ ID NO: 5465:  
US-09-107-532A-5465

Query Match 6.1%; Score 110.5; DB 4; Length 457;  
Best Local Similarity 23.7%; Pred. No. 0.0047;

Matches 59; Conservative 41; Mismatches 84; Indels 65; Gaps 13;  
QY 77 KLFIMET-PLFIASIVIGLITGLDNYLSYGLAYLPVSTSSLIIGTOLAFNALPA----- 130  
Db 12 KLILMTVPLLLIGNVFOQFYNMID-----MIIVGQTLGKNALAAVGATG 55  
QY 131 ---FLLVKQKFTPFPSINAVVLLTVGIGILALHSDGDKPAKESKEY-----VVGFLMTVVA 183  
Db 56 SLTFLI-----IGFAQGLTAGLAIITQRYGAKDYRGLKKSPAASVVISLVVTIL 106  
QY 184 ALL-YAFILPLVELTYKKARQEIFPLVLEIQVMCLAAATFFCVI--GMFIVGDFKVIAR 240  
Db 107 TVLSLLFIREMLQMQ-----TPPEILD-----QAQTFISIIILGIFASVFNLLSN 153  
QY 241 EAREPKIGGS---VFYVALIVITGIWQGFGLGAIGIVFCASSLASGLVLSVLLPV-TEV 296  
Db 154 VIR--ALGDSRTPLFLLIATVAVV-----LDLIFIYFGMGVEGNATVIAQVSSV 206  
QY 297 FAVVCFREK 305  
Db 207 LCLVYIKKK 215

Search completed: September 24, 2004, 00:02:07  
Job time : 24.6328 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:50:27 ; Search time 85.9831 Seconds  
(without alignments)  
1331.366 Million cell updates/sec

Title: US-09-913-767-8  
Perfect score: 1798  
Sequence: 1 MKNGLIINCIILITGTCGG.....DKPQPPETELPLPVSDYVA 356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1349238 segs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep:  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep:  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep:  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep:  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep:  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep:  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep:  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep:  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep:  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep:  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep:  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep:  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep:  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep:  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep:  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep:  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep:  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep:

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	956.5	53.2	320	12	US-10-424-599-173116 Sequence 173116,
2	931.5	51.8	306	12	US-10-424-599-216343 Sequence 216343,
3	882.5	49.1	293	12	US-10-425-114-44529 Sequence 44529, A
4	841.5	46.8	390	16	US-10-437-963-160022 Sequence 160022,
5	571.5	31.8	423	16	US-10-437-963-184544 Sequence 184544,
6	562.5	31.3	362	12	US-10-424-599-245361 Sequence 245361,
7	557	31.0	409	12	US-10-424-599-239885 Sequence 239885,
8	556.5	31.0	391	16	US-10-437-963-193908 Sequence 193908,
9	545	30.3	267	12	US-10-425-114-61519 Sequence 61519, A
10	535	29.8	751	16	US-10-437-963-146343 Sequence 146343,
11	532	29.6	372	16	US-10-437-963-136993 Sequence 136993,
12	531.5	29.6	355	12	US-10-424-599-174206 Sequence 174206,
13	521.5	29.0	366	12	US-10-424-599-145366 Sequence 145366,
14	516	28.7	363	12	US-10-424-599-174205 Sequence 174205,
15	514	28.6	379	12	US-10-425-114-45872 Sequence 45872, A

16	510	28.4	373	16	US-10-437-963-144907 Sequence 144907,
17	507	28.2	400	16	US-10-437-963-104380 Sequence 104380,
18	493.5	27.4	216	16	US-10-437-963-174935 Sequence 174935,
19	481	26.8	379	16	US-10-437-963-121341 Sequence 121341,
20	478.5	26.6	246	12	US-10-425-114-70607 Sequence 70607, A
21	474.5	26.4	327	12	US-10-424-599-218003 Sequence 218003,
22	465.5	25.9	610	16	US-10-437-963-186383 Sequence 186383,
23	403.5	22.4	238	12	US-10-424-599-204152 Sequence 204152,
24	324	18.0	241	12	US-10-424-599-152055 Sequence 152055,
25	272	15.1	202	16	US-10-767-701-38370 Sequence 38370, A
26	271.5	15.1	164	16	US-10-767-701-40564 Sequence 40564, A
27	246.5	13.7	173	12	US-10-424-599-257773 Sequence 257773,
28	229.5	12.8	166	16	US-10-767-701-35609 Sequence 35609, A
29	223	12.4	70	12	US-10-424-599-261293 Sequence 261293,
30	214	11.9	167	16	US-10-767-701-35590 Sequence 35590, A
31	184	10.2	150	16	US-10-767-701-37550 Sequence 37550, A
32	155	8.6	94	12	US-10-424-599-263994 Sequence 263994,
33	136.5	7.6	399	12	US-10-424-599-255573 Sequence 255573,
34	135	7.5	644	16	US-10-437-963-105606 Sequence 105606,
35	131.5	7.3	339	12	US-10-425-114-64888 Sequence 64888, A
36	131	7.3	358	16	US-10-437-963-186107 Sequence 186107,
37	130	7.2	382	12	US-10-425-114-62533 Sequence 62533, A
38	128.5	7.1	328	9	US-09-942-446-4 Sequence 4, Appli
39	128.5	7.1	328	9	US-09-795-693-35 Sequence 35, Appli
40	128.5	7.1	328	9	US-09-938-970-4 Sequence 4, Appli
41	128.5	7.1	328	14	US-10-156-239-35 Sequence 35, Appli
42	128.5	7.1	328	14	US-10-199-485-35 Sequence 35, Appli
43	128.5	7.1	328	14	US-10-252-646-7 Sequence 7, Appli
44	128.5	7.1	328	16	US-10-038-854-90 Sequence 90, Appli
45	127	7.1	358	16	US-10-437-963-198415 Sequence 198415,

ALIGNMENTS

RESULT 1

US-10-424-599-173116  
; Sequence 173116, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 173116  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12733C.1.pap  
US-10-424-599-173116

Query Match	53.2%	Score	956.5;	DB 12;	Length	320;
Best Local Similarity	57.0%;	Pred. No.	1.7e-86;			
Matches	184;	Conservative	62;	Mismatches	66;	Gaps 3;
QY	25	RLYFTNGKRIWFMSELSSTAGFPILIIPLLVSLSRGNNPNNNAENKRTKFLMETP	84			
Db	2	RLYFHGGKRIWLSSEFLTAANSARVLPITIIYHKKRRHS-----SALISIKPP	52			
QY	85	LFTASVIGLLTGLDNLVSYGLAVLPVSTSSLIIGTQAFNALFAFLVKQKTFPSIN	144			
Db	53	LFLASALIGLLGLDDYIACGVARLPVFTSLIQASHLAFTAVFAFLVHRFTYSVN	112			
QY	145	AVLLTVGIGILALHSDGPKAKESKKEYVGVGMTVAALLYAFILPLVELYKKARQ	204			
Db	113	SVLLTVAAVLALRSSGDRPAGESRSQVIGFVGMILAAAALYGFVPLMELVYKSRQ	172			

```

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44529
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700941791_FLI.pep
US-10-425-114-44529

Query Match          49.1%; Score 882.5; DB 12; Length 293;
Best Local Similarity 56.9%; Pred. No. 3.4e-79;
Matches 169; Conservative 60; Mismatches 57; Indels 11; Gaps 3;

Qy 51 IPLVSLRRRRGNPNNAENKRTKFLFMEPTPLFIASIVIGLLTGLDNYLSYGLAYL 110
Db 1 LPIIISVIHRRRHS-----SALISIKPPLFLASALIGLLTGLDNYLACGVARL 51

Qy 111 PVSTSSLIIGTQAFNALFAFLAVKQKFTPFSTINAVLLTVGIGILALHSDGDKPAKESK 170
Db 52 PVSTFSLIQASHLAFTAVFAFLVHRHFTPTYSVNSVLLTVAAVAVLALRSSGDRPAGES 111

Qy 171 KEYVVGFLMTVAALLYAFILPLVEITYKKARQETTFPLVLEIQMVNCLATFFCVIGMF 230
Db 112 RQVIGFVMTLAAALYGFVLPMLMELVYKSRQITVSLVMEIQVLCPFTLCTVGM 171

Qy 231 IVGDFKVIAREAREFKIGSVFYVALIVITGIIWQGFILGAIIVFCASSLASGLISVL 290
Db 172 INNDFKVIPEARDFKL-GETKYVVLVWSAIMWQFFLGAIGVIFCASSLLSGIIIAAF 230

Qy 291 LPVTEFVAVVCFREKFOAKGVSLLSLWGVSVFYGFESKGVKVDKPPPETELP 347
Db 231 LPVTEVLAVIVYKESFHAERGVALLVSLWGVSVFYGFEEKQDRE-KNKNRCPETDLP 286

RESULT 4
US-10-437-963-160022
; Sequence 160022, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 160022
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59340C.1.pep
US-10-437-963-160022

Query Match          46.8%; Score 841.5; DB 16; Length 390;
Best Local Similarity 48.0%; Pred. No. 5.9e-75;
Matches 169; Conservative 67; Mismatches 105; Indels 11; Gaps 4;

```

```

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44529
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700941791_FLI.pep
US-10-425-114-44529

Query Match          49.1%; Score 882.5; DB 12; Length 293;
Best Local Similarity 56.9%; Pred. No. 3.4e-79;
Matches 169; Conservative 60; Mismatches 57; Indels 11; Gaps 3;

Qy 51 IPLVSLRRRRGNPNNAENKRTKFLFMEPTPLFIASIVIGLLTGLDNYLSYGLAYL 110
Db 1 LPIIISVIHRRRHS-----SALISIKPPLFLASALIGLLTGLDNYLACGVARL 51

Qy 111 PVSTSSLIIGTQAFNALFAFLAVKQKFTPFSTINAVLLTVGIGILALHSDGDKPAKESK 170
Db 52 PVSTFSLIQASHLAFTAVFAFLVHRHFTPTYSVNSVLLTVAAVAVLALRSSGDRPAGES 111

Qy 171 KEYVVGFLMTVAALLYAFILPLVEITYKKARQETTFPLVLEIQMVNCLATFFCVIGMF 230
Db 112 RQVIGFVMTLAAALYGFVLPMLMELVYKSRQITVSLVMEIQVLCPFTLCTVGM 171

Qy 231 IVGDFKVIAREAREFKIGSVFYVALIVITGIIWQGFILGAIIVFCASSLASGLISVL 290
Db 172 INNDFKVIPEARDFKL-GETKYVVLVWSAIMWQFFLGAIGVIFCASSLLSGIIIAAF 230

Qy 291 LPVTEFVAVVCFREKFOAKGVSLLSLWGVSVFYGFESKGVKVDKPPPETELP 347
Db 231 LPVTEVLAVIVYKESFHAERGVALLVSLWGVSVFYGFEEKQDRE-KNKNRCPETDLP 286

RESULT 4
US-10-437-963-160022
; Sequence 160022, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 160022
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59340C.1.pep
US-10-437-963-160022

Query Match          46.8%; Score 841.5; DB 16; Length 390;
Best Local Similarity 48.0%; Pred. No. 5.9e-75;
Matches 169; Conservative 67; Mismatches 105; Indels 11; Gaps 4;

```

```

; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216343
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Glycine max
; NAME/KEY: unsure
; LOCATION: (1)..(306)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37387C.1.pep
US-10-424-599-216343

Query Match          51.8%; Score 931.5; DB 12; Length 306;
Best Local Similarity 62.6%; Pred. No. 4.8e-84;
Matches 186; Conservative 43; Mismatches 61; Indels 7; Gaps 3;

Qy 1 MKNGLIINCIIITGTCGPIILTRLYFTNGKRIWMFSLSTAGFPPIILPILVSLR 60
Db 15 MKRLITTCNLLTITGSGPLVMRLYFLHGGHVRVLSFLETAGFPLMLLPLAVSYLR 74

Qy 61 RRGNNPNNAENKRTKFLFMEPTPLFIASIVIGLLTGLDNYLSYGLAYLPVSTSLIG 120
Db 75 RR----TASAACTAKPKLISMKPPLAASTFTGILTGLDDYLYAGVARLPVSTSLIIA 130

Qy 121 TQAFNALFAFLAVKQKFTPFSTINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
Db 131 TQLGFTAFFAFLVLRQKFTAYSVNAVLLTVGVNGLALHTSGDRPGESVKYVWGVMT 190

Qy 181 VVAALYAFILPLVEITYKKARQETTFPLVLEIQMVNCLATFFCVIGMFIVGDFKVIAR 240
Db 191 VIAAALYGFILPLVELVYKIKQPLTVSLVMEIQVWCFSATFLCLLGMIIINNDFKVI 250

Qy 241 EAREFKIG-GSVFYVALIVITGIIWQGFILGAIIVFCASSLASGLISVLVPTVEV 296
Db 251 EAKKFEHGE--YVAVLNGSAILWQAFILGIVFWAMLLIXGLIIVLPVTEV 305

RESULT 3
US-10-425-114-44529
; Sequence 44529, Application US/10425114
; Publication No. US20040034889A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua

```





```
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 239885
; LENGTH: 409
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(409)
; OTHER INFORMATION: unsure at all xaa locations
; OTHER INFORMATION: Clone ID: PAT_MRT3847_58641C.1.pep
; US-10-424-599-239885

Query Match      31.0%; Score 557; DB 12; Length 409;
Best Local Similarity 37.1%; Pred. No. 1.3e-46;
Matches 122; Conservative 69; Mismatches 124; Indels 14; Gaps 3;

QY 5 LIIINCIILITIGCGPLLTRLYFTNGGKRIWFMFSLTAGPPIILPLLYSFLSRRGN 64
DB 66 LLIIVYACLVFVGSLSLSSLSKYFTFKGSSRWVSTWVOTAGPPLLLIPICVPILPKFT-K 124
QY 65 RNPNAENKRTKFLMETPFLFIASIVIGLLTGLDNLVSYGLAYLPVSTSLIIGTOLA 124
DB 125 RVFNDFTPR-----MLIISIGVMGFNNLFPFSGNSYLPVSTSAALLSSQLL 174
QY 125 FNALFAFLVKKQFTFPFSINAVLLTVGIGIILHSDGDKPAKESKEYVVGFLMTVVA 184
DB 175 ENLFSVILVKKITPSNVNVCVILLTSLIILHSDSSHERPKGLTKQNYFFIGFCTIGAG 234
QY 185 LLYAFILPLVELTYKKARQEIIFPLVLEIQVMVCLAAATFFCVIGMFVIGDFKVIAREARE 244
DB 235 LMFALYPLMEKIKYKKV---CYQVMEMQXIMEGATATAIIVGMTWDDGGFSEMKVSEQM 291
QY 245 FKIGSVFYVYALIVITGIWQGFGLGAGIVFCASSLASGVLSVLLPVTEVFVAVVCPRE 304
DB 292 VFDKGSRVYVTVMGNVTVWQLCFMGTAGVFLTSSLTGICWTFLLSNMVILGGVVFPRD 351
QY 305 KFOAEKGVLSLLSMGFVSFYGFGEKSGK 333
DB 352 AFGGVKAVSTFLCIGLFCSVYGYIKDNQ 380

RESULT 8
US-10-437-963-193908
; Sequence 193908, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barabazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 193908
; LENGTH: 391
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; NAME/KEY: unsure
```

```
; LOCATION: (1)..(391)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_90000C.1.pep
; US-10-437-963-193908

Query Match      31.0%; Score 556.5; DB 16; Length 391;
Best Local Similarity 37.1%; Pred. No. 1.3e-46;
Matches 132; Conservative 62; Mismatches 125; Indels 37; Gaps 6;

QY 5 LIIINCIILITIGCGPLLTRLYFTNGGKRIWFMFSLTAGPPIILPLLYS-FLSRRRG 63
DB 49 LVIFSACLVLG-ACGPELLLRVYFVHGTRLMLSATLQISGWPLLLPCLVSYRGRRHG 107
QY 64 RNPNAENKRTKFLMETPFLFIASIVIGLLTGLDNLVSYGLAYLPVSTSLIIGTQL 123
DB 108 IGN-----LLPRLVGAALVGLGVAVSCFVYALGSQLPLSTSLLLATQL 155
QY 124 AFNALFAFLVKKQFTFPFSINAVLLTVGIGIILHSDGDKPAKESKEYVVGFLMTVV- 182
DB 156 AFTAVFAFLVGLRFTFPFSANAVLLTVRVRGRCKTVRDPHGRGWHLDVQVVTLL 215
QY 183 -----AALYAPILPLVELTYKKARQEIIFPLVLEIQVMVCLAAATFFCVIGMFV 232
DB 216 QSQDWMMECGGVLGSELPP-----TSPYATVMQMAVMGAADTVXCVLIGMAIK 264
QY 233 GDFKVIAREAREFKIGSVFYVYALIVITGIWQGFGLGAGIVFCASSLASGVLSVLLP 292
DB 265 GAFQAVAREAAAFGL-GAANYVILVADAVSWQLNLGIMGLITCASSLLAGIMIAVLLP 323
QY 293 VTEVFVAVVCFREKFOAEKGVLSLLSMGFVSFYGFGEKSGKVKVVDKPPPETELPI 348
DB 324 LSQVLAVIFLHEKFDGKIGIALVLSLWGFASLYGERAKKQKGAQKMREREQVAL 379

RESULT 9
US-10-425-114-61519
; Sequence 61519, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; FILE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61519
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-024-D8_FLI.pep
; US-10-425-114-61519

Query Match      30.3%; Score 545; DB 12; Length 267;
Best Local Similarity 44.0%; Pred. No. 1.1e-45;
Matches 107; Conservative 55; Mismatches 67; Indels 14; Gaps 3;

QY 1 MKNGIILINCILITIGTCGPLLTRLYFTNGGKRIWFMFSLTAGPPIILPLLYSFLSR 60
DB 26 MRLIVATNCVWMLAGVTGGQLLSLYFKGHRQWLSGWLQTGCGWPLLLPVAASYVR 85
QY 61 REGNRPNNAENKRTKFLMETP--LFIASIVIGLLTGLDNLVSYGLAYLPVSTSLI 118
DB 86 RARDRS-----APVLLTQTQPRILLAAAGLGIAGVDNLLYAWGLEFLPVTSAI 136
QY 119 IGTQLAFNALFAFLVKKQFTFPFSINAVLLTVGIGIILHSDGDKPAKESKEYVVGFL 178
```

Db 137 ISQLATVLFAYLIVRQRUTMATVNAVALLTVGAVVLGHRVSSDRPAGVTRSQYWLGET 196  
Qy 179 MTVAALLYAFLPLVELTYKKAR---QEITFPLVLEIQVMVCLAAATFFCVIGFVIGDF 235  
Db 197 LTLGAALVYGLFLPLVELTYKCAAGGRPTVYALVVELQVLMGFVATVCTVGVNKNDF 256  
Qy 236 KVI 238  
Db 257 QVM 259

## RESULT 10

US-10-437-963-146343  
; Sequence 146343, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 146343  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_46977C.1.pap  
US-10-437-963-146343

Query Match 29.8%; Score 535; DB 16; Length 751;  
Best Local Similarity 33.9%; Pred. No. 4.3e-44;  
Matches 112; Conservative 74; Mismatches 130; Indels 14; Gaps 3;  
Qy 5 LIINCIIITIGTCGGPLRLTYFTNGGKRIWFMFLSTAGFPPIILPLVLSFLSRRGN 64  
Db 49 MVAVDVFFLIAGQTSATLLGRYYTQGRSKWISAFVRTAGFPILPFTLFF-----PS 102  
Qy 65 RNPNAENKRTKFLMETPLFIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIGTOLA 124  
Db 103 KSPSSCTNTPMAK-----LAVIYIVLGLIIAADDMMYTGGKLYLPASTYSLICASQLA 155  
Qy 125 FNALFAPLLVKQKFTPPSINAVLLTVGIGILALHSDGDKPAKESKEYVYVGFMTVVA 184  
Db 156 FNVVFSVLNSQKVTPLFNSVLLTMSASLIGSKESQGVTVGSGKYLILGFLVLTGAS 215  
Qy 185 LLYAFILPLVELTYKKARQBITFPLVLEIQVMVCLAAATFFCVIGMFTVIGDFKVIARE 244  
Db 216 CTYSLILALMQLFETIHKHTSFAVLNMQIYALVATAASVGLFASGEMWSLGEEMNA 275  
Qy 245 FKIGGSVFFYVALIITGIWQGFPLGAIGIVFCASSLASGLISVLPLVTEVFAVVC 304  
Db 276 FR-SGQFSYLMTLTLLMAVSNQVANIIGVLGIFEVSAFNSVISTVSLVPVFPFAVVVFD 334  
Qy 305 KFOAEKGVSLLSLWGFVSFYGFEPKSGKK 334  
Db 335 RMNGVKIVAMLIAWGFISYLFQHYLDGKK 364

## RESULT 11

US-10-437-963-136993  
; Sequence 136993, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:

; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; CURRENT FILING DATE: 2003-05-14  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 136993  
; LENGTH: 372  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_38519C.1.pap  
US-10-437-963-136993

Query Match 29.6%; Score 532; DB 16; Length 372;  
Best Local Similarity 36.8%; Pred. No. 3.4e-44;  
Matches 123; Conservative 67; Mismatches 122; Indels 22; Gaps 7;  
Qy 5 LIINCIIITIGTCGGPLRLTYFTNGGKRIWFMFLSTAGFPPIILPLVLSFLSRRGN 64  
Db 21 LFPANYAALLGGSVASSLSRYFAHGHNRWATLVQSVGFLLVPVY-----AG 72  
Qy 65 RNPNAENKRTKFLMET-PLFIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIGTOL 123  
Db 73 RSPS-----QPRPFAWFTRRLLAACVIGLVGMVGNLLFSYSSYLPVSTSSLIIGTOL 126  
Qy 124 AFNALFAPLLVKQKFTPPSINAVLLTVGIGILALHS--DGDPAKESKEYVYVGFMTV 181  
Db 127 AFTLVLAIIIVRHPLNFSNLNAVLLTSSVLIALRSSDSGEQPGGSRARYVGFVAVTL 186  
Qy 182 VAALLYAFILPLVELTYKKARQBITFPLVLEIQVMVCLAAATFFCVIGMFTVIGDFKVIARE 241  
Db 187 GAAGLFAAYLPVVELYVRKAVSG-GPRMAVEVQVIMQAATAALAVAGLAAGGWK---EE 242  
Qy 242 AREFKIGSVFFYVALIITGIWQGFPLGAIGIVFCASSLASGLISVLPLVTEVFAVVC 301  
Db 243 LARWDLSPAA-YWAVLAALVATWQCLMGTAGVYLTSSLSHSCVCMVTAULTANVIGVVV 301  
Qy 302 FREKFOAEKGVSLLSLWGFVSFYGFEPKSGKKV 335  
Db 302 FRDPFGADKAVATVLCVWGFSSYLYGEYTTQKKV 335

## RESULT 12

US-10-424-599-174206  
; Sequence 174206, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 174206  
; LENGTH: 355  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_128327C.1.pap  
US-10-424-599-174206

```

Query Match      29.6%; Score 531.5; DB 12; Length 355;
Best Local Similarity 35.0%; Pred. No. 3.6e-44;
Matches 115; Conservative 73; Mismatches 120; Indels 21; Gaps 4;

QY 5 LIIINCIILTTGTCGGPILLTRLYFTNGGKRIWFMWSFLSTAGFP---IILPLIVSFLSRR 61
DB 42 LLALSILAMLVAFPPASSILSRVYDNGGQSWIISWAVAGWPLTALILFP--VYFIS-- 97

QY 62 RGNRPNNAENKRTKLFMLMETPLFIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIGT 121
DB 98 -----KTFPTSLNLKLSLVIVGLFSAADNLMYAYAYLPASTASLVASS 144

QY 122 QIAFNALFAFLVKKQFTFFSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMTV 181
DB 145 SLVFSALFGYFLVKKNVNASIVNSVFIIITIALDSSDRYANISDSEYINGFVWDV 204

QY 182 VAALLYAFILPLVELTYKARQETFFPLVLEIQMWCLAAATFFCVIGMFIVGDFKVIARE 241
DB 205 LGSALHGLIFALSELVFKLERRSFIVLEQQVMVSLFAFLFTTVGMIMSGDFQGMARE 264

QY 242 AREFKIGGSVFYALIVITIGIQQGFFLGAIGVFCASSLASGVLISVLLPVTVEFVAVC 301
DB 265 ATTFK-GGRSAYLVIWGAITFQLGVLGGTAVIFLGSVTLVAGVLNAVTRTBITSAAIL 323

QY 302 PREKFAEKGVSLLSLMGFVSFYGFYGEFK 330
DB 324 LKDPMSGFKILSLVITFWGFGSYIYGSSK 352

RESULT 13
US-10-424-599-145366
; Sequence 145366, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 145366
; LENGTH: 366
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(363)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_102285C.1.pap
US-10-424-599-145366
Query Match      29.0%; Score 521.5; DB 12; Length 366;
Best Local Similarity 35.9%; Pred. No. 3.7e-43;
Matches 123; Conservative 66; Mismatches 133; Indels 21; Gaps 7;

QY 18 CGGP--LLTRLYFTNGGKRIWFMWSFLSTAGFPILILPLIVSFLSRRRGNRPNNAENK 75
DB 28 CGSAAILLGRLYEKGKSKWMTLVQLAGFP---IQLPFFISASK-NLTTNSSIHPKO 83

QY 76 TKLFLMETPLFIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIGTQLAFNALFAFLV 135
DB 84 SSASI-----LAFIVSGLLLALDCVLYSVGLWDLVPSTYSLICSSQLAFNAFFSYFLNS 139

QY 136 QKTFPFSINAVLLTVGIGILALHS-----GDKPAKESKKEYVVGFLMTVVAALLYA 188
DB 140 LKFTPIIINSLVLTISLTLVFNQSSDDDDSDSTKISKKKYIGFTICVAGSAGYG 199

QY 189 FILPLVELTYKARQETFFPLVLEIQMWCLAAATFFCVIGMFIVGDFKVIAREAFKIG 248
DB 200 LWLSLTQLVFKKVIKRETFKVVLDMLIYTSLVATLVLGVPASGEWSGLKNEKMEYEL- 258

QY 249 GSVFVYVALIVITIGIQQGFFLGAIGVFCASSLASGVLISVLLPVTVEFVAVVCFREKFOA 308
DB 259 GRASYLLNLTFTAILMQVFTIGCLGDISRVSSLSFSAISALGVPVPMPLAVLFFHDKMDG 318

QY 309 EKGVSLLLSLMGFVSFYGFYGEFKSKKVDK---POPPETELPI 348
DB 319 IKGISMVLAIMGIVSYVYQQYLLDDTKSENRRNTTSHVFKASSPI 361

RESULT 14
US-10-424-599-174205
; Sequence 174205, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 174205
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(363)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_128326C.1.pap
US-10-424-599-174205
Query Match      28.7%; Score 516; DB 12; Length 363;
Best Local Similarity 34.2%; Pred. No. 1.3e-42;
Matches 114; Conservative 73; Mismatches 124; Indels 22; Gaps 5;

QY 5 LIIINCIILTTGTCGGPILLTRLYFTNGGKRIWFMWSFLSTAGFP---IILPLIVSFLSRR 61
DB 48 LLVLSIILAMLVAFPPASSILSRVYDNGGQSWIISWAVAGWPLTALILFP--VYFIS-- 103

QY 62 RGNRPNNAENKRTKLFMLMETPLFIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIGT 121
DB 104 -----KTFPTPLNLKLSLVIVGLFSAADNLMYAYAYLPASTASLVASS 150

QY 122 QIAFNALFAFLVKKQFTFFSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMTV 181
DB 151 SLVFSALFGYFLVKKNVNASIVNSVFIIITIALTIIDSSDRYPSISDSEYINGFVWDV 210

QY 182 VAALLYAFILPLVELTYKARQETFFPLVLEIQMWCLAAATFFCVIGMFIVGDFKVIARE 241
DB 211 XGSAPHGLIFALSELVFKLGRSFIVLEQQVMVSLFAFLFTTVGMIVSGDFQGMARE 270

QY 242 AREFKIGGSVFYALIVITIGIQQGFFLGAIGVFCASSLASGVLISVLLPVTVEFVAVC 301
DB 271 ATTFK-GGRSAYLVIWGAITFQLGVLGGTAVIFLGSVTLVAGVLNAVTRTBITSAAIL 329

QY 302 PREKFAEKGVSLLSLMGFVSFYGFYGEFKSKK 334
DB 330 LKDPMSGFKILSLVITFWGFGSYIYGSSK 361

RESULT 15
US-10-425-114-45872
; Sequence 45872, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
```

Search completed: September 24, 2004, 00:06:41  
Job time : 87.9831 secs

**THIS PAGE BLANK (UPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:40:51 ; Search time 15.0847 Seconds  
(without alignments)  
1228.856 Million cell updates/sec

Title: US-09-913-767-8

Perfect score: 1798

Sequence: 1 MKNGLIINCIIITICGCG.....DKQPPELPILPVS DYVA 356

Scoring table: BLOSUM62

Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	144	8.0	311	1 YCAY_CLOKL	P38943 clostridium
2	141	7.8	470	1 PTTE_BACSU	P39794 bacillus su
3	134.5	7.5	343	1 Y841_METTH	O26929 methanobact
4	124	6.9	287	1 YXCE_BACSU	P40420 bacillus su
5	123	6.8	364	1 FCT1_HUMAN	Q96a29 homo sapien
6	121	6.7	629	1 CTR1_HUMAN	P30825 homo sapien
7	120.5	6.7	389	1 YDHP_ECOLI	P77399 escherichia
8	119	6.6	286	1 YXCE_BACME	P40419 bacillus me
9	118	6.6	479	1 PTSE_VIBAL	P22825 vibrio algi
10	116.5	6.5	389	1 YDHP_ECO57	Q8x625 escherichia
11	114	6.3	556	1 NU2M_PODAN	P15578 podospora a
12	113	6.3	759	1 NAHB_ONCMY	Q01345 oncorhynchu
13	111	6.2	532	1 YC18_HAEIN	Q57251 haemophilus
14	110.5	6.1	330	1 YETK_BACSU	O31540 bacillus su
15	110.5	6.1	547	1 Y447_MYCGE	P47685 mycoplasma
16	109.5	6.1	512	1 EMRY_ECOLI	P52600 escherichia
17	109	6.1	349	1 PERM_HABIN	P43969 haemophilus
18	108.5	6.0	342	1 YE44_YEAST	P40004 saccharomyc
19	108	6.0	445	1 YPNF_BACSU	P54181 bacillus su
20	108	6.0	480	1 PTSE_STAXY	P51184 staphylococ
21	107	6.0	353	1 GMS1_SCHPO	P87041 schizosacch
22	107	6.0	393	1 NUPC_BACSU	P39141 bacillus su
23	106.5	5.9	532	1 COX1_RHOCA	P98059 rhodobacter
24	106	5.9	324	1 PSTC_XYLFA	Q9pbk2 xyella fas
25	105.5	5.9	305	1 YVBV_BACSU	O32256 bacillus su
26	105.5	5.9	378	1 YF69_AQUAE	O67513 aquifex aeo
27	105.5	5.9	615	1 UAPA_EMENI	Q07307 emericeila
28	105	5.8	485	1 YIHO_ECOLI	P32136 escherichia
29	104.5	5.8	394	1 YWFF_BACSU	P39642 bacillus su
30	104.5	5.8	511	1 XASA_ECO57	P58229 escherichia
31	104.5	5.8	511	1 XASA_ECOLI	P39183 escherichia
32	104	5.8	268	1 UPKI_RHILO	Q98dm7 rhizobium l
33	104	5.8	325	1 MRAY_BACHD	Q9k986 bacillus ha

#### ALIGNMENTS

RESULT 1

YCAY\_CLOKL STANDARD; PRT; 311 AA.

AC P38943;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical transport protein in catl 5'region (ORFY).  
OS Clostridium kluyveri.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1534;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 555;  
RX MEDLINE=96146540; PubMed=8550525;  
RA Soehling B., Gottschalk G.;  
RT "Molecular analysis of the anaerobic succinate degradation pathway in  
Clostridium kluyveri.";  
RL J. Bacteriol. 178:871-880(1996).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -!- SIMILARITY: Belongs to the eamA transporter family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
between the Swiss Institute of Bioinformatics and the EMBL outstation -  
the European Bioinformatics Institute. There are no restrictions on its  
use by non-profit institutions as long as its content is in no way  
modified and this statement is not removed. Usage by and for commercial  
entities requires a license agreement (See http://www.isb-sib.ch/announce/  
or send an email to license@isb-sib.ch).  
CC -----  
CC EMBL; L21902; AAA92345.1; -  
DR InterPro; IPR000620; DUF6.  
DR Pfam; PF00892; DUF6; 2.  
KW Hypothetical protein; Transport; Transmembrane.  
FT TRANSMEM 6 26 POTENTIAL.  
FT TRANSMEM 33 53 POTENTIAL.  
FT TRANSMEM 70 90 POTENTIAL.  
FT TRANSMEM 97 117 POTENTIAL.  
FT TRANSMEM 123 143 POTENTIAL.  
FT TRANSMEM 155 175 POTENTIAL.  
FT TRANSMEM 185 205 POTENTIAL.  
FT TRANSMEM 219 239 POTENTIAL.  
FT TRANSMEM 244 264 POTENTIAL.  
FT TRANSMEM 265 285 POTENTIAL.  
SQ SEQUENCE 311 AA; 33507 MW; A5E7CB9EC9ADF0B CRC64;

Query Match 8.0%; Score 144; DB 1; Length 311;

Best Local Similarity 19.1%; Pred. No. 0.0032;

Matches 73; Conservative 68; Mismatches 113; Indels 128; Gaps 17;

Qy 1 MKNGLIINCIIITICGCGPLLTLYFTN---GGK-----RIWFKSFLSTAGF 46

Db 1 MKGYIFI-----LTAIFYSTQEISGRMLAQKQAMDFQWMIYFLIGA-- 45

Qy 47 PIILPLLVSLRRRRGNPNNAENKRRTKL-----FLMETPLFIASIVIGLLTGLDN 100

Db 46 -IILLPAVKDI-----KVKKLKLTGNDLGYLALCGILAVSISMSML----- 86

Qy 101 YLYSYGLAYLPVSTSSLIIGTQAFNALFAPLLVKKQKFTFPSINAVLLTVGIGILALHS 160

Db 87 ---OFAVTYTKASTAAVLFCNTAVFTTIPFAYFILKEKIKGITIVSIIVSLIGVWIF--- 140

Qy 161 DGDPAKESK---KEYVVGFLMTWAAALYAFILFELVELTYKARQETFPFLVLEIQMV 216

Db 141 ---NPAKMEGIGSRDLIGCFALVAVWSLYTVI-----SKKRIEIVGVFN--- 198

Qy 217 MCLAAATPCVIGMFTVGVDFKVIAREAREFKTGGSVFYVALIVITG-----I 262

Db 189 -CI-SFFPGVIALI-----LUVVTGRPIFSGITLNNILVL 222

Qy 263 IWQGFPLGAIGIVFC-----ASSLASGVLSVLLPVTVEFVAVVCFRKFQAEKGVSL 315

Db 223 LYMGIFIKAVGYI-CYLGAIKETSAVTASTVFLIKPALATVLAIIIGESIEVNVVIGIV 281

Qy 316 LSLWGFVSFYGEFKSGKKVD 337

Db 282 FIIIGSII-----NYSNKKAND 299

RESULT 2

PTTB BACSU

ID PTTB BACSU STANDARD; PRT; 470 AA.

AC P39794; 034771;

DT 01-FEB-1995 (Rel. 31, Created)

DT 30-MAY-2000 (Rel. 39, Last annotation update)

DT 15-MAR-2004 (Rel. 43, Last annotation update)

DE Probable PTS system, trehalose-specific IIBC component (EIBC-TRE)

DE (Trehalose-permease IIBC component) (Phosphotransferase enzyme II, BC component) (EC 2.7.1.69) (EII-TRE).

GN TREP OR TREB OR BSU07800.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423;

RN [1]\_SEQUENCE FROM N.A.

RP STRAIN=168 / Marburg;

RC MEDLINE=97074649; PubMed=8917076;

RX Schoeck F., Dahl M.K.;

RA "Analysis of DNA flanking the treA gene of Bacillus subtilis reveals genes encoding a putative specific enzyme iITre and a potential regulator of the trehalose operon."

RT Gene 175:59-63 (1996).

RL [2]

RN [3]

RP SEQUENCE FROM N.A.

RC STRAIN=AC327;

RX MEDLINE=97124190; PubMed=8969503;

RA Yamamoto H., Uchiyama S., Sekiguchi J.;

RT "Cloning and sequencing of a 40.6 kb segment in the 73 degrees-76 degrees region of the Bacillus subtilis chromosome containing genes for trehalose metabolism and acetoin utilization."

RT Microbiology 142:3057-3065 (1996).

RN [4]

RP SEQUENCE FROM N.A.

RC STRAIN=AC327;

RX MEDLINE=97417488; PubMed=9272861;

RA Yamamoto H., Uchiyama S., Nugroho F.A., Sekiguchi J.;

RT "Cloning and sequencing of a 35.7 kb in the 70 degrees-73 degree region of the Bacillus subtilis genome reveal genes for a new two-component system, three spore germination proteins, an iron uptake system and a general stress response protein."

RT Gene 194:191-199 (1997).

RL [4]

RN SEQUENCE FROM N.A.

RP STRAIN=168;

RC MEDLINE=98044033; PubMed=9384377;

RX Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Besieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Conterton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Chim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G., Guisepi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Klaer-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kunano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadale Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scofield F., Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.;

RT "The complete genome sequence of the Gram-positive bacterium Bacillus subtilis."

RL Nature 390:249-256 (1997).

RN [5]

RP SEQUENCE OF 324-470 FROM N.A.

RC STRAIN=168;

RX MEDLINE=95379486; PubMed=7651129;

RA Helfert C., Gotsche S., Dahl M.K.;

RT "Cleavage of trehalose-phosphate in Bacillus subtilis is catalysed by a phospho-alpha-(1-1)-glucosidase encoded by the treA gene."

RL Mol. Microbiol. 16:111-120 (1995).

CC -1- FUNCTION: This is a component of the phosphoenolpyruvate-dependent sugar phosphotransferase system (PTS), a major carbohydrate active -transport system. The IICD domains contain the sugar binding site and the transmembrane channel; the IIA domain contains the primary phosphorylation site (the donor is phospho-HPr); IIA transfers its phosphoryl group to the IIB domain which finally transfers it to the sugar.

CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein histidine + sugar phosphate.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.

CC -1- SIMILARITY: Contains 1 PTS EIIIB domain.

CC -1- SIMILARITY: Contains 1 PTS EIIIC domain.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@sib-sib.ch](mailto:license@sib-sib.ch)).

CC -----

CC EMBL; Z54245; CAA91014.1; -

DR EMBL; D83967; BAA23409.1; -

DR EMBL; D86417; BAA22289.1; -

DR EMBL; Z99108; CAB12609.1; -

DR EMBL; X80203; CAA56494.1; -

DR PIR; C69725; C69725.

DR HSSP; P05053; 1IBA.

DR Subtilist; BG11009; trep.

DR InterPro; IPR001396; Pfam EIIIB.

DR InterPro; IPR003352; Pfam EIIIC.

DR InterPro; IPR004719; PHSIIC GIC.

DR Pfam; PF00367; PTS EIIIB; 1.

DR Pfam; PF02378; PTS EIIIC; 1.

DR ProDom; PD001476; PTS EIIIB; 1.

DR TIGRFAMs; TIGR00826; EIIIB gic; 1.

DR TIGRFAMs; TIGR00852; pts-Glc; 1.



DR PROSITE; PS01035; PTS\_EIIB\_CVS; 1.  
 KW Phosphotransferase system; Sugar transport; Transferase;  
 KW Transmembrane; Inner membrane; Phosphorylation; Complete proteome.  
 FT DOMAIN 1 ? EIIB.  
 FT MOD\_RES 26 26 PHOSPHORYLATION (BY SIMILARITY).  
 FT TRANSMEM 110 130 POTENTIAL.  
 FT TRANSMEM 160 180 POTENTIAL.  
 FT TRANSMEM 183 203 POTENTIAL.  
 FT TRANSMEM 234 254 POTENTIAL.  
 FT TRANSMEM 263 283 POTENTIAL.  
 FT TRANSMEM 301 321 POTENTIAL.  
 FT TRANSMEM 326 346 POTENTIAL.  
 FT TRANSMEM 347 367 POTENTIAL.  
 FT TRANSMEM 375 395 POTENTIAL.  
 FT TRANSMEM 403 423 POTENTIAL.  
 FT TRANSMEM 443 463 POTENTIAL.  
 FT CONFLICT 140 140 F -> S (IN REF. 1).  
 FT CONFLICT 363 363 M -> L (IN REF. 1 AND 5).  
 FT CONFLICT 465 465 A -> G (IN REF. 1 AND 5).  
 SQ SEQUENCE 470 AA; 49999 MW; 7A741850A2697D53 CRC64;  
 Query Match 7.8%; Score 141; DB 1; Length 470;  
 Best Local Similarity 23.5%; Pred. No. 0.0073;  
 Matches 87; Conservative 59; Mismatches 112; Indels 112; Gaps 20;  
 QY 64 NRPNNNAENKRTKFLMETPTPIASIVIGLLTGLDNLVSYGLAYLPVSTSS-----116  
 Db 99 NNPPLQRAVKTLDIFI--PILPAIVTAGLLGINNLTAEIGFF--STKSIQVYYPQ 152  
 QY 117 -----LIIGTQLAF-NALPAFLVVKQFPFSSINAVLLTVGIGIALHSD-----161  
 Db 153 WADLANMINLIAGTAFPLALIGSAVKR-----FGGNP--LLGIVLGMVLPVHLLNAW 206  
 QY 162 GDKPAKESKKEYV-----VGLMTVVAALLVAFILPLVELTVKKARQE-ITFPLV 210  
 Db 207 GYGAARSQSEIIPWNI-FGLEVKQVGGQVLPILLASIMLAKLEVLTFRTPGIIQLLV 266  
 QY 211 LEIQVMCLAAATPCVIG--MFIQVDPF---KVIAREAREPKIGSVF--YYALIVITGI- 262  
 Db 267 APITLLTGFASPI-IIGPTTFAIGNVLTSLISVFGSFAALGGLLYGGPYSALVITGMH 325  
 QY 263 -----IWQGFPL-----GAIGIVPCAS-----SLASGVLSV 289  
 Db 326 HTFLAVLDLQIGKGGTFLWPMALSNIAQGSAAALAMFIVRDEKQKGLSLTSG--ISA 383  
 QY 290 LLPVTR--VFAY--VCPREF-----QAEKGYSLLSLWGFVSFYF 326  
 Db 384 YLGITPAIFGVNLRYPPIIAMVSSGLAGMYISSQGVLLASVGGVGPPIFSIMSYW 443  
 QY 327 GEFKSGKKV 336  
 Db 444 GAFAIGMAIV 453  
 RESULT 3  
 Y841 METH STANDARD; PRT; 343 AA.  
 AC Q26929;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical transport protein MTH841.  
 GN MTH841.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,

RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lum W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwani N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 J. Bacteriol. 179:7135-7155(1997).  
 RL -SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -SIMILARITY: Belongs to the eamA transporter family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; AE000861; AAB85339.1; -  
 DR PIR; C69212; C69212.  
 DR InterPro; IPR000620; DUF6.  
 DR Pfam; PF00892; DUF6; 2.  
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.  
 FT TRANSMEM 13 33 POTENTIAL.  
 FT TRANSMEM 44 64 POTENTIAL.  
 FT TRANSMEM 71 91 POTENTIAL.  
 FT TRANSMEM 121 141 POTENTIAL.  
 FT TRANSMEM 148 168 POTENTIAL.  
 FT TRANSMEM 177 197 POTENTIAL.  
 FT TRANSMEM 203 223 POTENTIAL.  
 FT TRANSMEM 244 264 POTENTIAL.  
 FT TRANSMEM 269 289 POTENTIAL.  
 FT TRANSMEM 296 316 POTENTIAL.  
 FT TRANSMEM 320 340 POTENTIAL.  
 SQ SEQUENCE 343 AA; 37839 MW; 4296C7890CB91B7F CRC64;  
 Query Match 7.5%; Score 134.5; DB 1; Length 343;  
 Best Local Similarity 21.3%; Pred. No. 0.016;  
 Matches 77; Conservative 62; Mismatches 130; Indels 93; Gaps 17;  
 QY 14 TIGTCGGPLLRLYFTNGG--KRIW-FMGFLSTAGF-----PI-----48  
 Db 28 TISMCGCP-----BEGAYMRRLMGYSITITATIFFGISATLKDMLSSMHPVTIGAYT 80  
 QY 49 -ILIPLLVGSFLS-----RRGNRPNNNAENKRTKFLMETPLFIASIVIGLL 95  
 Db 81 YIAGIFLPFIREPTLRHVNAINRKGE---SEARIKRDYLLILLTALL--STVIAPL 135  
 QY 96 TGLDNLVSYGLAYLPVSTSSLIIGTQLAFNALPAFLVKKQKTFPSINAVLLTVGIGI 155  
 Db 136 -----LPLTGLGDDTAVNASLIINVEVLFILLYLIPRETQLQKDFLIGVILII--LGA 187  
 QY 156 LALHSDGDKPAKESKKEYVVGFLMTVVAALLYAFILPLVELTVKKARQETFFLVLEIQM 215  
 Db 188 VYLLTEGDF-STILRNNAVTVGNFLVMAAFAFFWS--LDTVLSKFLSKRDLIF-----ISG 239  
 QY 216 VMCLAAATPCVIGMFVGDVKVIAREAREPKIGSVYFVALIVITGIIWGGFLGAIG--273  
 Db 240 VKSSVGGFVLLIIMLILGINTLEPLMLPYALGVSVF-----SIGCS 281  
 QY 274 --IVFCASSLASGVLSVLLPVTE---VFVVCFREKFOAEKGVSLLSLWGFVSFYF 326  
 Db 282 FLDIYIATREIGASVMGALFPLSLFGAIFAFILREPFISIMQGISIVMLTGVFIIDYN 341  
 QY 327 GE 328  
 Db 342 GK 343  
 RESULT 4  
 YCXE\_BACS

CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC -----  
 CC ENBL; D50453; BAA09023.1; --  
 CC ENBL; Z39106; CAB12200.1; --  
 CC ENBL; M23547; -; NOT\_ANNOTATED\_CDS.  
 CC ENBL; M12276; -; NOT\_ANNOTATED\_CDS.  
 CC PIR; G69766; G69766.  
 CC Subtilisin; BG11066; ycxX.  
 CC InterPro; IPR004673; Rhat; 1.  
 CC TIGRFAMS; TIGR00776; Rhat; 1.  
 CC Hypothetical protein; Sporulation; Complete proteome.  
 KW SEQUENCE 287 AA; 30892 MW; E0FF3B0B47EBC56A CRC64;  
 SQ  
 Query Match 6.9%; Score 124; DB 1; Length 287;  
 Best Local Similarity 20.7%; Pred. No. 0.07;  
 Matches 57; Conservative 45; Mismatches 87; Indels 86; Gaps 11;  
 QY 79 FLMETELFTASIVIGLLTGLDNLVSYGLA-----YLPVSTSLIIGTQLAF 125  
 DB 47 FFVQPVLSLRFITVIGVSGL---FWSLQANQLKSLQMLGVSKTMTPIST-----GMQLVS 98  
 QY 126 NALFAFLVKKQKFTPSIN---AVLLTVGIGILALHSDGDKPAKES---KKEYVVGFL 178  
 DB 99 TSLGVVFVFEWSTPIATLGVALLFIIVGILTSLEKDKKKEGSPNLKK---GIL 154  
 QY 179 MTVAALLYAFILVELTYLTKARQBITFPLVLEIQMVLMCLATFFCVIGMFVIGDKVI 238  
 DB 155 ILLVSTLGY-----LVVVVARLFNVGWSAL 181  
 QY 239 AREAREKIGSGVPE-----YVALIVITGIWOG--PFL---GAIGVFCASSLAS 283  
 DB 182 LPQAGVMVVGGLVTVRHKFNKATRNILPLGIWAGNMLFISQPRGVGATFSLSQM 241  
 QY 284 GVLISVLLPVTVEFVAVVCFREKFAEKGVSLLSL 318  
 DB 242 GIVISTL-----GGIFILREKTKRQLIAIAGI 270  
 RESULT 5  
 PCT1 HUMAN STANDARD; PRT; 364 AA.  
 ID FCT1 HUMAN STANDARD; PRT; 364 AA.  
 AC Q96A29; Q9B776; Q9NUJ8;  
 DT 28-FEB-2003 (Rel. 41, Created)  
 DT 28-FEB-2003 (Rel. 41, Last sequence update)  
 DT 15-MAR-2004 (Rel. 43, Last annotation update)  
 DE GDP-fucose transporter 1.  
 GN FUCT1.  
 OS Homo sapiens (Human).  
 OC Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;  
 OC Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.  
 OX NCBI\_TaxID=9606;  
 RN [1]  
 RP SEQUENCE FROM N.A., CHARACTERIZATION, AND VARIANT CDG-IIC CVS-147.  
 RX MEDLINE=21225289; PubMed=11326279;  
 RA Luehn K., Wild M.K., Eckhardt M., Gerardy-Schahn R., Vestweber D.;  
 RT "The gene defective in leukocyte adhesion deficiency II encodes a  
 RT putative GDP-fucose transporter.";  
 RL Nat. Genet. 28:69-72(2001).  
 RN [2]  
 RP SEQUENCE FROM N.A., AND VARIANTS CDG-IIC CVS-147 AND ARG-308.  
 RX MEDLINE=21225290; PubMed=11326280;  
 RA Luecke T., Marquardt T., Etzioni A., Hartmann E., von Figura K.,  
 RA Koerner C.;  
 RT "Complementation cloning identifies CDG-IIC, a new type of congenital  
 RT disorders of glycosylation, as a GDP-fucose transporter deficiency.";  
 RL Nat. Genet. 28:73-76(2001).  
 RN [3]  
 RP SEQUENCE FROM N.A.  
 RN TISSUE=Mammary gland, and Placenta;

YCXE BACSU STANDARD; PRT; 287 AA.  
 P40420; P94429;  
 01-FEB-1995 (Rel. 31, Created)  
 30-MAY-2000 (Rel. 35, Last sequence update)  
 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Hypothetical protein ycxX.  
 GN YCXE OR BSU03920.  
 OS Bacillus subtilis.  
 OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
 OX NCBI\_TaxID=1423;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=97124189; PubMed=8969502;  
 RA Yanane K., Kumano M., Kurita K.;  
 RT "The 25 degrees-36 degrees region of the Bacillus subtilis chromosome:  
 RT determination of the sequence of a 146 kb segment and identification  
 RT of 113 genes.";  
 RL Microbiology 142:3047-3056(1996).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=168;  
 RX MEDLINE=98044033; PubMed=9384377;  
 RA Kunat F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
 RA Azevedo V., Bortner M.G., Bessieres P., Bolotin A., Borchert S.,  
 RA Boriss R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
 RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
 RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
 RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emerson P.T.,  
 RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
 RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Gallier N.,  
 RA Ghim S.Y., Glaser P., Goffeau A., Gollightly E.J., Grandi G.,  
 RA Giuseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
 RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
 RA Joris B., Karamata D., Kasahara Y., Klaerr-Blanchard M., Klein C.,  
 RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
 RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
 RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,  
 RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
 RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
 RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
 RA Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
 RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
 RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
 RA Sekiguchi J., Sekowska A., Seror S.J., Serror P., Shin B.S., Soldo B.,  
 RA Sorokin A., Taccani E., Takagi T., Takahashi H., Takemaru K.,  
 RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
 RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
 RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,  
 RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
 RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
 RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
 RT subtilis.";  
 RL Nature 390:249-256(1997).  
 RN [3]  
 RP SEQUENCE OF 1-9 FROM N.A.  
 RX MEDLINE=89033889; PubMed=3141376;  
 RA Rather P.N., Moran C.P. Jr.;  
 RT "Compartment-specific transcription in Bacillus subtilis:  
 RT identification of the promoter for gdh.";  
 RL J. Bacteriol. 170:5086-5092(1988).  
 RN [4]  
 RP SEQUENCE OF 272-287 FROM N.A.  
 RX MEDLINE=86168021; PubMed=3082854;  
 RA Lampel K.A., Uratani B., Chaudhry G.R., Ramaley R.F., Rudikoff S.;  
 RT "Characterization of the developmentally regulated Bacillus subtilis  
 RT glucose dehydrogenase gene.";  
 RL J. Bacteriol. 166:238-243(1986).  
 CC -1- DEVELOPMENTAL STAGE: Expressed during sporulation (By similarity).  
 CC -1- SIMILARITY: TO A SIMILAR ORF IN B.MEGATERIUM.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation



-!- SUBCELLULAR LOCATION: Integral membrane protein.  
-!- TISSUE SPECIFICITY: Ubiquitous.  
-!- SIMILARITY: BELONGS TO A FAMILY THAT GROUPS MANY AMINO ACID PERMEASES.

-----  
This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See http://www.isb-sib.ch/announce/ or send an email to license@isb-sib.ch).

-----  
EMBL; X59155; CAA41869.1; --  
EMBL; X57303; CAA40560.1; --  
EMBL; AF078107; AAC27721.1; --  
PIR; S29685; S29685.  
Genew; HGNC:11057; SLC7A1.  
MIM; 104615; --  
GO; GO:0005887; C:integral to plasma membrane; TAS.  
InterPro; IPR002293; AA/rel\_permease1.  
InterPro; IPR004755; Cat\_AA\_permease.  
InterPro; IPR004841; Permease\_region.  
Pfam; PF00324; aa\_permeases; 2.  
TIGRFAMs; TIGR00906; 2A0303; 1.  
Transmembrane; Glycoprotein; Transport; Amino-acid transport;  
Receptor.

DOMAIN	1	35
TRANSMEM	36	57
DOMAIN	58	61
TRANSMEM	62	82
DOMAIN	83	102
TRANSMEM	103	123
DOMAIN	124	162
TRANSMEM	163	183
DOMAIN	184	191
TRANSMEM	192	212
DOMAIN	213	246
TRANSMEM	247	267
DOMAIN	268	287
TRANSMEM	288	307
DOMAIN	308	337
TRANSMEM	338	358
DOMAIN	359	384
TRANSMEM	385	405
DOMAIN	406	408
TRANSMEM	409	429
DOMAIN	430	432
TRANSMEM	433	513
DOMAIN	514	526
TRANSMEM	527	551
DOMAIN	552	559
TRANSMEM	560	580
DOMAIN	581	584
TRANSMEM	585	605
DOMAIN	606	629
CARBOHYD	226	226
CARBOHYD	235	235
CONFLICT	23	23
SEQUENCE	629 AA;	67638 MW; 717734D4/93647C5 CRC64;

Query Match  
Best Local Similarity 23.6%; Pred. No. 0.22;  
Matches 72; Conservative 47; Mismatches 112; Indels 74; Gaps 16;

53 LLVSPLSRGRGNPNNAENKTKFLMETPLFTASIVIGLLTGDLNLYSYGLAVLPV 112  
|||:::||||:  
370 LLEFKFLA-----NVNDRTKTPT-----IATLASGAVAAMAFLED-----LKD 407  
|||:::||||:  
113 STSLSLGIQTALNFALFALLKQKTPFSINAVLLTVGIGILAHSDGDKPAKE----- 168  
|||:::||||:  
408 LVDMISIGITLLAYSLVAACVLV-LRVQPQPNLVQO-----MASTSELDPADONELA 459  
|||:::||||:

Query Db Qy

DR	InterPro: IPR005828; Sub_transporter.	
DR	FrAm; PF00083; sugar tr; 1.	
DR	PROSITE; PS00850; MFS; 1.	
KW	Hypothetical protein; Transport; Transmembrane; Inner membrane;	
KW	Complete proteome.	
FT	TRANSMEM 7	POTENTIAL.
FT	TRANSMEM 44	POTENTIAL.
FT	TRANSMEM 64	POTENTIAL.
FT	TRANSMEM 71	POTENTIAL.
FT	TRANSMEM 91	POTENTIAL.
FT	TRANSMEM 101	POTENTIAL.
FT	TRANSMEM 121	POTENTIAL.
FT	TRANSMEM 131	POTENTIAL.
FT	TRANSMEM 151	POTENTIAL.
FT	TRANSMEM 160	POTENTIAL.
FT	TRANSMEM 180	POTENTIAL.
FT	TRANSMEM 204	POTENTIAL.
FT	TRANSMEM 224	POTENTIAL.
FT	TRANSMEM 237	POTENTIAL.
FT	TRANSMEM 257	POTENTIAL.
FT	TRANSMEM 267	POTENTIAL.
FT	TRANSMEM 287	POTENTIAL.
FT	TRANSMEM 291	POTENTIAL.
FT	TRANSMEM 311	POTENTIAL.
FT	TRANSMEM 331	POTENTIAL.
FT	TRANSMEM 357	POTENTIAL.
FT	TRANSMEM 377	POTENTIAL.
SQ	SEQUENCE 389 AA; 40064 MW; 90A6A22CBD565BF3 CRC64;	

  

Query Match	6.7%; Score 120.5; DB 1; Length 389;
Best Local Similarity	25.1%; Pred. No. 0.16;
Matches	74; Conservative 42; Mismatches 96; Indels 83; Gaps 14;

  

QY	13	LTIGTCGG-PLLNRLYFTNGGKRWFMFSLSSTAGFPILLPLLYSVFLSRRGNGRPNNAE	71
DB	138	LTIANIGGVPAATWLGSTIG-----WRMSFLATAGLVISMVSL--FFSLPKGGA--GAR	188
QY	72	NKERTKLFLEMETPLFIASIVIGLL-TGLDNLYSYGLAYLPVSTSSLLIIGQLAFNALFA	130
DB	189	PEVKKELAVLMRRQVLSALLTVLGAGAMTLTYISFVLSQITHA-----	234
QY	131	FLIVKQKTFPSINAVLLTVGIGILALHSDGDKPAKESKEYVVGFLMTVVAALLYAFI	190
DB	235	-----TPVFVTAMLVL-IGVFSIGNVLGGLADRSVNGTLKGFLLLLMWIML--A	282
QY	191	LPLVELTYKARQETFPVLVLEITQMWCLATPCV-----IG	228
DB	283	IPFL-----ARNEFG-----AASIMVVMGAAATFAVPFLOQRMVRVASEAPGLSSSVNIG	332
QY	229	MFTVGDGPKVIARAREPKIGSVFY----YALIVITGIIWQGFPLGAGIVFCAS	279
DB	333	AFNLGN-----ALGAAGAVISAGLGSFVPMGAIVAGL---ALLLVFMSA	377

  

RESULT 8		
YCKE	BACME	
ID	YCKE	BACME
AC	STANDARD;	PRT; 286 AA.
AC	P40419;	
DT	01-FEB-1995	(Rel. 31, Created)
DT	01-FEB-1995	(Rel. 31, Last sequence update)
DT	16-OCT-2001	(Rel. 40, Last annotation update)
DE	Hypothetical 30.5 kDa protein in gdhI 5' region (ORF 2).	
OS	Bacillus megaterium.	
OC	Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.	
OX	NCBI_TaxID=1404;	
RN	[1]	
RC	SEQUENCE FROM N.A.	
RP	STRAIN=IAM 1030;	
RA	Mitamura T., Ebora R.V., Nakai T., Makino Y., Negoro S., Urabe I.,	
RA	Okada H.;	
RT	"Structure of isozyme genes of glucose dehydrogenase from Bacillus	
RL	megaterium IAM1030."	
RL	J. Ferment. Bioeng. 70:363-369(1990).	
CC	-1- DEVELOPMENTAL STAGE: Expressed during sporulation.	
CC	-1- SIMILARITY: TO A SIMILAR ORF IN B.SUBTILIS.	
CC	-----	
CC	This SWISS-PROT entry is copyright. It is produced through a collaboration	
CC	between the Swiss Institute of Bioinformatics and the EMBL Outstation -	
CC	the European Bioinformatics Institute. There are no restrictions on its	
CC	use by non-profit institutions as long as its content is in no way	
CC	modified and this statement is not removed. Usage by and for commercial	
CC	entities requires a license agreement (See <a href="http://www.isb-sib.ch/announce/">http://www.isb-sib.ch/</a> ).	
CC	or send an email to <a href="mailto:license@isb-sib.ch">license@isb-sib.ch</a> .	

```

CC EMBL; D90043; BAA14098.1; --
CC DR PIR; I39851; JS0384.
CC DR InterPro; IPR004673; Rhat.
CC DR TIGRFAMs; TIGR00776; Rhat; 1.
CC KW Hypothetical protein; Sporulation.
CC SQ SEQUENCE 286 AA; 30490 MW; 95AB89D0251ld74D CRC64;

Query Match 6.6%; Score 119; DB 1; Length 286;
Best Local Similarity 77.2%; Pred. No. 0.15; 79; Indels 56; Gaps 11;
Matches 63; Conservative 34; Mismatches 79;

QY 83 TPL-FTASIVIGLL--TGLDNYLSYGL-----AYLPVSTSLIIGTQAFNALFAFLVK 135
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 191 LPLFGVGVVSGLFWAQSNQLKSIDLIGVSKTPIST-----GLQLVSTSLFGVIVFH 108
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 136 QKTPPSI-----NAVULLTVGIGILAHSDGDKPAKESKEY-----VVGFLMTV 181
    : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : ||| : |||
Db 109 EWSTKTSIILGVNALIFIIVIGVLASLQSEKEAEKGKFNKGIIVILLISTVGYLVV 166
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 182 VAALLY-----AFILPL-----VELTVKK-----ARQETFPVL-----LRIQM 215
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 169 VVARLFNVDCWSALLPQAGVIGVLLTFKHPFNKYAIRNIPGLIAGNMFILSQ 228
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

QY 216 VMCLATFFCVGMFIV---GDPKVIAREAREFKIGSVFYVALIVITGII 263
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||
Db 229 PKGVATSFSLSQWIVISTLGLIILGEKTKKQLVGIIGIILIIAGVM 280
    ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| ||| |||

RESULT 9
PTSB VIBAL STANDARD; PRT; 479 AA.
ID PTSB VIBAL
AC P22825;
DT 01-AUG-1991 (Rel. 19, Created)
DT 01-AUG-1991 (Rel. 19, Last sequence update)
DT 15-MAR-2004 (Rel. 43, Last annotation update)
DE PTS system, sucrose-specific IIBC component (EIIBC-SCR) (Sucrose-
DE perase IIBC component) (Phosphotransferase enzyme II, BC component)
DE (EC 2.7.1.69) (EII-SCR).
GN SCRA.
OS Vibrio alginolyticus.
OC Bacteria; Proteobacteria; Gammaproteobacteria; Vibrionales;
OC Vibrionaceae; Vibrio.
OX NCBI_TaxID=563;
RN [1]
RP SEQUENCE FROM N.A.
RX MEDLINE=91071601; PubMed=2174811;
RA Blatch G.L., Scholle R.R., Woods D.R.;
RT "Nucleotide sequence and analysis of the Vibrio alginolyticus sucrose
RT uptake-encoding region.";
RL Gene 95:17-23(1990).
RN [2]
RP SEQUENCE OF 1-12 FROM N.A.
RX MEDLINE=91285433; PubMed=2060795;
RA Blatch G.L., Woods D.R.;
RT "Nucleotide sequence and analysis of the Vibrio alginolyticus scr
RT repressor-encoding gene (scrR).";
RL Gene 101:45-50(1991).
CC -1- FUNCTION: This is a component of the phosphoenolpyruvate-dependent
CC sugar phosphotransferase system (PTS), a major carbohydrate active
CC -transport system. The IICD domains contain the sugar binding site
CC and the transmembrane channel; the IIA domain contains the primary
CC phosphorylation site (the donor is phospho-HPr); IIA transfers its
CC phosphoryl group to the IIB domain which finally transfers it to
CC the sugar.
CC -1- CATALYTIC ACTIVITY: Protein N-phosphohistidine + sugar = protein
CC histidine + sugar phosphate.
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane.
CC -1- SIMILARITY: Contains 1 PTS EIIA domain.
CC -1- SIMILARITY: Contains 1 PTS EIIC domain.
CC This SWISS-PROT entry is copyright. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation

```

**BEST AVAILABLE COPY**

the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (see <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; M76768; AAA27555.1; -  
EMBL; M35009; AAA27557.2; -  
PIR; J00781; JQ0781.  
HSP; P05053; 11BA.  
InterPro; IPR001996; Ptrans EIIB.  
InterPro; IPR003352; Ptrans EIIC.  
InterPro; IPR004719; Ptrans EIIC.  
Pfam; PF00367; Ptrans EIIB; 1.  
Pfam; PF02378; Ptrans EIIC; 1.  
Pfam; PF001476; Ptrans EIIB; 1.  
TIGRFAMs; TIGR00826; EIIB-GIC; 1.  
TIGRFAMs; TIGR00852; Ptrans-GIC; 1.  
PROSITE; P01035; Ptrans EIIB-CYS; 1.  
Phosphotransferase system; Sugar transport; Transferase;  
Transmembrane; Inner membrane; Phosphorylation.  
DOMAIN 1 ? EIIB.  
DOMAIN ? 479  
MOD RES 26 26 PHOSPHORYLATION (BY SIMILARITY).  
MOD RES 324 324 PHOSPHORYLATION (BY SIMILARITY).  
SEQUENCE 479 AA; 49890 MW; FBF906B5170E3EB7 CRC64;

Query Match 6.6%; Score 118; DB 1; Length 479;  
Best Local Similarity 28.4%; Pred. No. 0.28; Indels 48; Gaps 11;  
Matches 62; Conservative 28; Mismatches 80; Indels 48; Gaps 11;  
84 PLFIASIVIGLTLGDNVLYSYGLAYLPVSTSSLIIG-----TQLA-----FNALFAPLL 133  
115 PIIPAIVAGLLMGVINGLITAGLGF---IDGKSLIEANPGLTDLANMINTFANAPFVLP 171  
134 VKQKFT---PFSINAVLLTVGIGIALHSD-----GDKP-----AKESKE 172  
172 ILLAFSASKKFGGNP---YLGAALGMLVHVPDLNMGFGGASVSGNIPWNILGFEIQK- 228  
173 YVGFMTVVVAALYAFILPLVELTYKKARQITPLVLEIQMVMCLAAATFFCVIGM--- 229  
229 --VYGQGVLPVLVSFAFLAKVELGKRVIPSVLDNLLTFLAIFLAGLITTVVGPTR 286  
230 ---FIVGDFKVIAREEFKIGSVF--YVALIVITGI 262  
287 DIGFLIGDGLNWLNTAGF-VGGAFLIYAPFVITGM 323

RESULT 10  
YDHP ECO57 STANDARD; PRT; 389 AA.  
AC ORX625;  
DT 28-FEB-2003 (Rel. 41, Created)  
DT 28-FEB-2003 (Rel. 41, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical transport protein ydhp.  
GN YDHP OR 22679 OR ECS2366.  
OS Escherichia coli O157:H7.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
OC Enterobacteriaceae; Escherichia.  
OK NCBI\_TaxID=83334;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=O157:H7 / EDL933 / ATCC 700927;  
RA MEDLINE=21074935; PubMed=11206551;  
RA Perna N.T., Plunkett G. III, Burland V., Mau B., Glasner J.D.,  
RA Rose D.J., Mayhew G.F., Evans P.S., Gregor J., Kirkpatrick H.A.,  
RA Posral G., Hackett J., Klink S., Boutin A., Shao Y., Miller L.,  
RA Grobeck E.J., Davis N.W., Lim A., Dimalanta E.T., Potamoudis K.,  
RA Apodaca J., Anantharaman T.S., Lin J., Yen G., Schwartz D.C.,  
RA Welch R.A., Blattner F.R.;  
RT "Genome sequence of enterohaemorrhagic Escherichia coli O157:H7";  
RL Nature 409:529-533(2001).

[2] SEQUENCE FROM N.A.  
RP STRAIN=O157:H7 / RIMD 0509952;  
RX MEDLINE=21156231; PubMed=11258796;  
RA Hayashi T., Makino K., Okuyama K., Kurokawa K., Ishii K., Yokoyama K.,  
RA Han C.-G., Ohtsubo E., Nakayama K., Murata T., Tanaka M., Tobe T.,  
RA Iida T., Takami H., Honda T., Sasaki C., Ogasawara N., Yasunaga T.,  
RA Kuhara S., Shiba T., Hattori M., Shinagawa H.,  
RA "Complete genome sequence of enterohaemorrhagic Escherichia coli  
RT O157:H7 and genomic comparison with a laboratory strain K-12";  
RL DNA Res. 8:11-22(2001).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (potential).  
CC -1- SIMILARITY: BELONGS TO THE MAJOR FACILITATOR FAMILY (ALSO KNOWN  
CC AS THE DRUG RESISTANCE TRANSLOCASE FAMILY).  
CC -----  
CC This SWISS-PROT entry is copyrighted. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

EMBL; AE005389; AAG56646.1; -  
EMBL; AP002558; BAB35789.1; -  
PIR; B85773; B85773.  
PIR; F90924; F90924.  
InterPro; IPR007114; MFS.  
InterPro; IPR005828; Sub\_transporter.  
Pfam; PF00083; sugar tr; 1.  
PROSITE; PS0850; MFS; 1.  
Hypothetical protein; Transport; Transmembrane; Inner membrane;  
Complete proteome. 27 POTENTIAL.  
FT TRANSMEM 7 27 POTENTIAL.  
FT TRANSMEM 44 64 POTENTIAL.  
FT TRANSMEM 71 91 POTENTIAL.  
FT TRANSMEM 101 121 POTENTIAL.  
FT TRANSMEM 131 151 POTENTIAL.  
FT TRANSMEM 160 180 POTENTIAL.  
FT TRANSMEM 204 224 POTENTIAL.  
FT TRANSMEM 237 257 POTENTIAL.  
FT TRANSMEM 267 287 POTENTIAL.  
FT TRANSMEM 291 311 POTENTIAL.  
FT TRANSMEM 331 351 POTENTIAL.  
FT TRANSMEM 357 377 POTENTIAL.  
SQ SEQUENCE 389 AA; 40035 MW; B53864327068881E CRC64;

Query Match 6.5%; Score 116.5; DB 1; Length 389;  
Best Local Similarity 26.1%; Pred. No. 0.3; Indels 67; Gaps 14;  
Matches 75; Conservative 43; Mismatches 102; Indels 67; Gaps 14;  
13 LTIGTCGG-PLLTLYFTNGSKRIWMSFLSTAGFPILIPLVSLFSLRRNRNPNNAE 71  
138 LTLANIGGVPAATWLGETIG-----WRMSFLATAGLVISVSL--FSLPKGGA---GAR 188  
72 NKRTKLFMETPLFIASIVIGL-TGLDNVLYSYGLAYLPVSTSSLIIGTQLAFNALFA 130  
189 PEVKKELAVLRPQVLSALLTVLGAAGMTLYTYISPLVQSITHA----- 234  
131 FLLVKQKFTPFSINAVLLTVGIGIALHSDGDKPAKESKEYVYVGFMTVVAAALLYAFI 190  
235 -----TPVFVTAMLV-LGVFSGNYLGGKLAORSVNGTLKGFL-----LLMVM 280  
191 LPLVELTYKKARQEI-----TFPLVLEIQM-VMCLAA-----TFPCVIGMETVGDFFK 236  
281 LAIPFLARNKFGAALSMVGAATFVVPPLQMRVMRVASEAPGLSSVNVIGAFNLGN-- 338  
237 VTARBARFKIGGSVFY-----YALIVITGIWQGFELGAIGVFCA 279  
339 -----ALGAAAGGAVISAGLGYSFVPMVAIVAGL---ALLVPM 377





CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC -----  
CC EMBL, U22801; AAC22871.1; --  
DR FIR; H64110; H64110.  
DR TIGR; H11218; --  
DR InterPro; IPR003804; Lactate perm.  
DR Pfam; PF02652; Lactate\_perm; 1.  
DR TIGRFAMs; TIGR00795; lctP; 1.  
DR Hypothetical protein; Transport; Transmembrane; Inner membrane;  
KW Complete proteome.  
FT TRANSMEM 23 43 POTENTIAL.  
FT TRANSMEM 56 76 POTENTIAL.  
FT TRANSMEM 101 121 POTENTIAL.  
FT TRANSMEM 129 149 POTENTIAL.  
FT TRANSMEM 152 172 POTENTIAL.  
FT TRANSMEM 180 200 POTENTIAL.  
FT TRANSMEM 213 233 POTENTIAL.  
FT TRANSMEM 234 254 POTENTIAL.  
FT TRANSMEM 274 294 POTENTIAL.  
FT TRANSMEM 346 366 POTENTIAL.  
FT TRANSMEM 387 407 POTENTIAL.  
FT TRANSMEM 420 440 POTENTIAL.  
FT TRANSMEM 462 482 POTENTIAL.  
FT TRANSMEM 508 528 POTENTIAL.  
SQ SEQUENCE 532 AA; 56884 MW; F3C1DF47E54217FB CRC64;  
  
Query Match 6.2%; Score 111; DB 1; Length 532;  
Best Local Similarity 21.3%; Pred. No. 0.95;  
Matches 78; Conservative 56; Mismatches 138; Indels 94; Gaps 15;  
  
QY 6 IINCIILITGTCGGPL-----LTRYFTNGKRIWFMFSLSTAGPPIILIPLLVSF 57  
DB 147 LIMNSVPVSGAVGTPTFWFGALKUSEMILIGSITAFIHSIAALIPLALRIIV-- 204  
  
QY 58 LSRRRGRNPNNAENKTKLFL-----METPLFIASIV-----IGLLTGLDN 100  
DB 205 -----NWDDIRKNIVFIVISVLGCVFVFLIAQVNYEFPSPVIGVGAIGFISVWA 253  
  
QY 101 YLYSYGLAYLPVSTSLIIGTQAFNALF-----AELLVKQ---KTFPFSINAVLLT 150  
DB 254 ANRNIGLAKVTNLDNNAVSAGEVVKALFPTGLLIAFLIVRIHQLPFKAMMDATWFS 313  
  
QY 151 VGIGILAL-----HSDGDKPAKSKKEYV---VGFLMTVVVAALLYAFILPVE 195  
DB 314 TTLGSLGLFEISKGLIFSLKNIFGNSVSSYKLLVYPALIPFVITVLA-----IPFFK 367  
  
QY 196 LTYKKARQEIITPVLBIQVMCLAAATFFCVIGMFIQDFKVIAREAREPKIGSVFYA 255  
DB 368 ISSNNVKQ---ILVSSLQK---SKNPFIALIGALVMVNLMLVGGEHSMVKIIGRTF--- 417  
  
QY 256 LIVTGIWQGF--FLGAIQIVFCASSLASGVLI--SVLLPVTVEFVAVVCFREKQAKGV 312  
DB 418 -AEISGNSNTIFFSFLGAIGSFFSGSNVTNLSITFGSVQLSTAETGI-----SV 465  
  
QY 313 SILLSL 318  
DB 466 ALVLAL 471  
  
RESULT 14  
YETK\_BACSU  
ID\_YETK\_BACSU STANDARD; PRT; 330 AA.  
AC Q31540;  
DT 15-JUL-1999 (Rel. 38, Created)  
DT 15-JUL-1999 (Rel. 38, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical transport protein yetk.  
GN YETK OR BSU07210.  
  
FT DOMAIN 331 352 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 353 372 M8 (POTENTIAL).  
FT DOMAIN 373 376 CYTOPLASMIC (POTENTIAL).  
FT TRANSMEM 377 398 M9 (POTENTIAL).  
FT DOMAIN 399 446 EXTRACELLULAR (POTENTIAL).  
FT TRANSMEM 447 467 M10 (POTENTIAL).  
FT DOMAIN 468 759 CYTOPLASMIC (POTENTIAL).  
FT MOD\_RES 641 641 PHOSPHORYLATION (BY PKA) (POTENTIAL).  
FT MOD\_RES 648 648 PHOSPHORYLATION (BY PKA) (POTENTIAL).  
FT CARBOHYD 49 49 N-LINKED (GLCNAC. .) (POTENTIAL).  
FT CARBOHYD 338 338 N-LINKED (GLCNAC. .) (POTENTIAL).  
SQ SEQUENCE 759 AA; 85173 MW; D6D95442995AE251 CRC64;  
  
Query Match 6.3%; Score 113; DB 1; Length 759;  
Best Local Similarity 19.5%; Pred. No. 0.94;  
Matches 57; Conservative 41; Mismatches 83; Indels 112; Gaps 11;  
  
QY 146 VLLTVGILAHSDGDKPAKSKKEYVGFMTVAALLYATILPLVELTYKKARQEI 205  
DB 17 VIVLVVFVGI-GLPIEASAPAYQS--HGTEGSHLTNTTKAF--PVLAVNVEHVRKPF 71  
  
QY 206 TFPVLVEIQMVNCLA-----ATPFCVIGMFIQDFKVIARE----- 241  
DB 72 EIALWILLALLMKUGFLHPLRSVAVPSCLLIVVGLLVGLGKIVIGEPFVLDLSQLFFL 131  
  
QY 242 -----AREPKIG-GSVFYVALIVITGIIWQGFILGAIIVFC----- 277  
DB 132 CLLPPIILDAGYFLPIRPPTENVGTILVFAVI---GTLWNAFFFMGLLYALCOIESVGLS 188  
  
QY 278 -----ASSLASGVLSVLLPVT----- 294  
DB 189 GVDLLACLFLGFSIVSAVDVAVLAVFEIHNELWHILVFGESLNDVATVVLNLPREF 248  
  
QY 295 -----EVR-AVVCPEKFOAKGVSLLSLMGFVSFYFGEFKSGKKVVD 337  
DB 249 SKVGTVTVLVDVLGVVCF---FVSLGGVLVGAIVGLAAFTSRFTSTRVIE 298  
  
RESULT 13  
YCI8\_HAEIN  
ID\_YCI8\_HAEIN STANDARD; PRT; 532 AA.  
AC Q57251;  
DT 01-NOV-1997 (Rel. 35, Created)  
DT 01-NOV-1997 (Rel. 35, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Putative L-lactate permease.  
GN H11218.  
OS Haemophilus influenzae.  
OC Bacteria; Proteobacteria; Gammaproteobacteria; Pasteurellales;  
OC Pasteurellaceae; Haemophilus.  
OX NCBI\_TaxID=727;  
RN (1)  
RP SEQUENCE FROM N.A.  
RC STRAIN=RD / KW20 / ATCC 51907;  
RX MEDLINE=95350630; PubMed=7542800;  
RA Fleischmann R.D., Adams M.D., White O., Clayton R.A., Kirkness E.F.,  
RA Kerlavage A.R., Bult C.J., Tomb J.-F., Dougherty B.A., Merrick J.M.,  
RA McKenney K., Sutton G., Fitzhugh W., Fields C.A., Gocayne J.D.,  
RA Scott J.D., Shirley R., Liu L.-I., Glodek A., Kelley J.M.,  
RA Weidman J.F., Phillips C.A., Spriggs T., Hedblom E., Cotton M.D.,  
RA Utterback T.R., Hanna M.C., Nguyen D.T., Saudek D.M., Brandon R.C.,  
RA Fine L.D., Fritchman J.L., Fuhrmann J.L., Geoghegan N.S.M.,  
RA Gnehm C.L., McDonald L.A., Small K.V., Fraser C.M., Smith H.O.,  
RA Venter J.C.;  
RT "Whole-genome random sequencing and assembly of Haemophilus influenzae  
RT Rd.";  
RL Science 269:496-512(1995).  
CC -!- FUNCTION: May play a role in L-lactate transport.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane  
CC (Potential).  
CC -!- SIMILARITY: Belongs to the lactate permease family.  
CC -----  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration



QY	77 KLFMETPUP-----IASIVIGLL-----TGIDNYLYS 104	FT TRANSMEM	53	POTENTIAL.
	:          : :	FT TRANSMEM	107	POTENTIAL.
	:          : :	FT TRANSMEM	145	POTENTIAL.

Mon Sep 27 08:51:15 2004

```
FT TRANSMEM 203 223 POTENTIAL.
FT TRANSMEM 231 251 POTENTIAL.
FT TRANSMEM 263 283 POTENTIAL.
FT TRANSMEM 298 318 POTENTIAL.
FT TRANSMEM 351 371 POTENTIAL.
FT TRANSMEM 397 417 POTENTIAL.
FT TRANSMEM 432 452 POTENTIAL.
FT TRANSMEM 470 490 POTENTIAL.
FT TRANSMEM 499 519 POTENTIAL.
SQ SEQUENCE 547 2A; 62052 MW; E6D115C419A2E81B CRC64;

Query Match 6.1%; Score 110.5; DB 1; Length 547;
Best Local Similarity 19.8%; Pred. No. 1.1;
Matches 73; Conservative 64; Mismatches 140; Indels 91; Gaps 16;

QY 5 LIINCIILITGTGCGPLLTRLYFTNGGKRI---WFMSFLSTAGFPFIILPLLVFLSR- 60
Db 109 IVLLNAINIFPLGTGVIFSKAIGNDQNKIQEAWNTGLISTTVFGLITQPLVLSPFAKEW 168
QY 61 RRGNRNPNNAENK---RKTCLFLMETPLFIASIVIGLTLGDNLYLSYGLAYLPVSTSSL 117
Db 169 LHYNLDQSSFEQNFQANSFQOFENKKAIDVASEVYVILIGL-NIIPMLSRLLFFYLAQSE- 226
QY 118 IIGTOL-----AFNALFAPLLVKKQKTPFPSINAVLLTVGIGILALHSDGDKPEAK 167
Db 227 --GRQLFAIVPPIANLINILIVFLLVRY-----SLGVIGSAVAG----- 265
QY 168 ESKKEYVVGFMTVVAALLYAFIL---PLVELTYKKAR-QBITFPLVLEIQMVNCLAAATF 223
Db 266 -----ILGYLINFLAYIILYLNKRNLTLYTKIKLNKIDFNLLVVVSL----- 311
QY 224 FCVIGM---FIVGDFPKVIAREAREPKIG-----GSVFYVALIVITGIWQGFILGAI 272
Db 312 ---IGWASFFRNGSLSIVTTFYESFLVNLTKATTDKNDVFYLTLLT-----GPI 357
QY 273 GIVFCASSIASGVLSVLLPVTEFVAVVCFREKFOAEKGVSLLSL-----WGFVSYFYGE 328
Db 358 AI-----SNLASAAIFGLLQGVRTVSSYKFGQKKYDEIKINIYTVIICISFGSLIYLTA 413
QY 329 FKSQKKV 336
Db 414 VAFGKQIL 421
```

Search completed: September 23, 2004, 23:50:21  
Job time : 19.0847 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:42:01 ; Search time 24.8588 Seconds  
(without alignments)  
1362.072 Million cell updates/sec

Title: US-09-913-767-9  
Perfect score: 1760  
Sequence: 1 MKALVINCIIILAIICNGCG.....QQEESQETSLSRPISSEC 352  
Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues  
Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	1744.5	99.1	351	2 C86408	F3H9.12 protein -
2	1149	65.3	356	2 B84749	hypothetical prote
3	1128	64.1	356	2 D86408	purine permease [i
4	557.5	31.7	344	2 T04924	hypothetical prote
5	557	31.6	1128	2 T04922	hypothetical prote
6	556.5	31.6	379	2 D96506	hypothetical prote
7	555.5	31.6	348	2 T04923	hypothetical prote
8	530	30.1	315	2 A84634	hypothetical prote
9	502.5	28.6	358	2 T04921	hypothetical prote
10	475	27.0	432	2 C85087	hypothetical prote
11	357	20.3	393	2 D86330	hypothetical prote
12	282	16.0	390	2 D96513	unknown protein T1
13	279.5	15.9	392	2 H96516	F16N3.10 [imported
14	277.5	15.8	389	2 E96516	F16N3.13 [imported
15	262.5	14.9	398	2 C96613	hypothetical prote
16	261.5	14.9	383	2 H86232	hypothetical prote
17	140.5	8.0	323	2 D84597	hypothetical prote
18	134.5	7.6	368	2 T10557	hypothetical prote
19	134	7.6	461	2 B97305	probable cation ef
20	134	7.6	518	2 I39989	spore cortex synth
21	133.5	7.6	345	2 AC1252	conserved hypothet
22	131.5	7.5	287	2 E75159	hypothetical prote
23	130	7.4	292	2 D71181	hypothetical prote
24	129	7.3	296	2 B83384	conserved hypothet
25	129	7.3	311	2 F83698	hypothetical prote
26	129	7.3	345	2 AI1614	conserved hypothet
27	127.5	7.2	288	2 F72219	conserved hypothet
28	127	7.2	298	2 A83892	hypothetical prote
29	127	7.2	325	2 D75555	conserved hypothet

conserved hypothet  
NADH2 dehydrogenas  
MadN protein VC007  
conserved hypothet  
sodium-dependent t  
licB protein limpo  
hypothetical prote  
permease [imported  
probable phosphate  
gluconate permease  
hypothetical prote  
conserved hypothet  
hypothetical prote  
C4-dicarboxylate t  
NADH2 dehydrogenas  
NADH2 dehydrogenas

ALIGNMENTS

RESULT 1

C86408

F3H9.12 protein - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 27-Nov-2001

C;Accession: C86408

R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Pederspiel, N.A.; Kaul, S.; White, O.; Alonso,

Chin, C.W.; Hughes, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;

ansen, N.F.; Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.;

Nature 408, 816-820, 2000

A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, B.; Kim, C.;

C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.

ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A;Reference number: A86141; MUID:21016719; PMID:11130712

A;Accession: C86408

A;Status: preliminary

A;Molecule type: DNA

A;Residues: 1-351 <STO>

A;Cross-references: GB:AE005172; NID:g9795615; PIDN:AAF98433.1; GSPDB:GN00141

C;Genetics:

A;Map position: 1

C;Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 99.1%; Score 1744.5; DB 2; Length 351;  
Best Local Similarity 99.7%; Pred. No. 3e-125;  
Matches 351; Conservative 0; Mismatches 0; Indels 1; Gaps 1;

QY	1	MKALVINCIIILAIICNGCGPLIMRLYPNNGKRIWFSTFLETAGFPVIFIPILFSYITR	60
DB	1	MKALVINCIIILAIICNGCGPLIMRLYPNNGKRIWFSTFLETAGFPVIFIPILFSYITR	60
QY	61	RSNNVGDSTSPFLIKPRLLIAAVIVGILSGFDNYLYAYGIVLPYSTALIIASOLAFI	120
DB	61	RSNNVGDSTSPFLIKPRLLIAAVIVGILSGFDNYLYAYGIVLPYSTALIIASOLAFI	120
QY	121	AIFSPFMVKHKTPTFTINAVLLTVGAAVLGMHTETDKPVHETHKQYITGFLITVA	180
DB	121	AIFSPFMVKHKTPTFTINAVLLTVGAAVLGMHTETDKPVHETHKQYITGFLITVA	180
QY	181	YAFILPLVELAYQKAKQTMSTYLVLEFQILICLLASIVSGIMFVIGAGDFKALPKGAREF	240
DB	181	YAFILPLVELAYQKAKQTMSTYLVLEFQILICLLASIVSGIMFVIGAGDFKALPKGAREF	240
QY	241	KLGEALFYVAVFSALIIWQGFILGALIFSTSSLVSGIMISVLLPITVLAIVFVHEKF	300
DB	240	KLGEALFYVAVFSALIIWQGFILGALIFSTSSLVSGIMISVLLPITVLAIVFVHEKF	299
QY	301	QAEKGLSLALSUNGFVSFYFGIKSGEDKRRIQEESQETSLSRPISSEC	352
DB	301	QAEKGLSLALSUNGFVSFYFGIKSGEDKRRIQEESQETSLSRPISSEC	352

Db 300 QAERGLSLALSLWGFVSFYGEIKSGEDKRRIOQESQETESLSRPISEC 351

RESULT 2

B84749

hypothetical protein At2g33750 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: B84749

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: B84749

A>Status: Preliminary

A:Molecule type: DNA

A:Residues: 1-356 <STO>

A:Cross-references: GB:AE002093; NID:g1707019; PIDN:AAC69140.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g33750

A:Map position: 2

C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 65.3%; Score 1149; DB 2; Length 356;  
Best Local Similarity 60.4%; Pred. No. 5.1e-80;  
Matches 218; Conservative 68; Mismatches 57; Indels 18; Gaps 3;

QY 1 MYKALVIINCIIILAIAGNCGGLIMRLYFNNGKRIWFSTFLETAGFPVIFIPLLFSYITR 60  
DB 1 MKTVLIVINCIIILAIAGNCGGLIMRLYFNNGKRIWFSTFLETAGFPVIFIPLLFSYITR 60

QY 61 RRNNVGDSTSPFLIKPRLLIAAIVIGILSGFDNLYAYGAYLVPSTAALIIASQLAPI 120  
DB 61 RCLSEQETTPFLMKPPLFAIAIVGLLVGFDNLYAYGAYLVPSTAALIIASQLGFT 120

QY 121 AIPSPFMVKHKTPTFTINAVLLTGAAVLGMHTETDKPVHETHKQYITGFLITVAAVM 180  
DB 121 ALFAFMVKQKFTPTFTINAVLLTGAAVLGMHTETDKPVHETHKQYITGFLITVAALL 180

QY 181 YAFILPLVELAYOKAKOTMSYTLVLEFQILCLLASIVSVIGMFIAGDFK----- 230  
DB 181 YGFILPLVELSKYKSGRIYITALEFQVNLFAATCVCLVGMKLAAGDFKVKHALFIFKN 240

QY 231 QALPKAREFKLGEALFYVAVFSAIIWQGFPLGALIGLIFSTSSLVSGIMISVLLPITEV 290  
DB 241 RVIAEGARDFKLGSLLYVIVFTAIWQAFVGAIGLIFCASSILVSGIMVSALLPVTVI 300

QY 291 LAVIFHEKFOAEKGLSLALSLSLWGFVSFYGEIKSGEDKRRIOQESQETESLSRPI 350  
DB 301 LAVICFOEKFOAKGVALALSLSLWGSVSFYGVQVKSEE-----KTKAQDTQSUL--PVT 352

QY 351 E 351  
DB 353 D 353

RESULT 3

D86408

purine permease [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: D86408

R:Theologas, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D86408

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-356 <STO>

A:Cross-references: GB:AE005172; NID:g9795614; PIDN:AAF98432.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 64.1%; Score 1128; DB 2; Length 356;  
Best Local Similarity 60.1%; Pred. No. 2e-78;  
Matches 215; Conservative 66; Mismatches 65; Indels 12; Gaps 5;

QY 1 MYKALVIINCIIILAIAGNCGGLIMRLYFNNGKRIWFSTFLETAGFPVIFIPLLFSYITR 60  
DB 1 MKNGLIINCIIILAIAGNCGGLIMRLYFNNGKRIWFSTFLETAGFPVIFIPLLFSYITR 60

QY 61 RRNNVGDSTSPFLIKPRLLIAAIVIGILSGFDNLYAYGAYLVPSTAALIIA 114  
DB 61 RRNRNPNNNAENKRTKFLMETPLFIASIVIGLLTGLDNLVSYGLAYLPVSTSSLIIG 120

QY 115 SOLAFIAIPSPFMVKHKTPTFTINAVLLTGAAVLGMHTETDKPVHETHKQYITGFLIT 174  
DB 121 TQLAFNALFAFLVQKQFTFPFSINAVLLTGVIGILALHSDGDKPAKESKKEYVGVFLMT 180

QY 175 VAAAVMYAFTPLPLVELAYOKAKOTMSYTLVLEFQILCLLASIVSVIGMFIAGDFKQALP 234  
DB 181 VWAALLYAFILPLVELYTKARQETFPPLVLEIQMWCLAAATFCVIGMFIAGDFK-VIA 239

QY 235 KEAREFKLGEALFYVAVFSAIIWQGFPLGALIGLIFSTSSLVSGIMISVLLPITEVLA 293  
DB 240 REAREFKLGSVSFYALIVITGIWQGFPLGALIGIVFCASSLASGLVLSVLLPVTVEFAV 299

QY 294 IFYHEKFOAEKGLSLALSLSLWGFVSFYGEIKSGEDKRRIOQESQETESLSRPISE 351  
DB 300 VCFREKFOAEKGVLSLLSLWGFVSFYGEFKSG--KKVWDKQPQPPETELPIL--PVS 353

RESULT 4

T04924

hypothetical protein T9A21.70 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Jun-2000

C:Accession: T04924

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X. submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15390

A:Accession: T04924

A:Molecule type: DNA

A:Residues: 1-344 <BEV>

A:Cross-references: EMBL:AL021713

A:Experimental source: cultivar Columbia; BAC clone T9A21

C:Genetics:

A:Map position: 4

A:Introns: 11/1

A>Note: T9A21.70

C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 31.7%; Score 557.5; DB 2; Length 344;  
Best Local Similarity 35.3%; Pred. No. 4.1e-35;  
Matches 120; Conservative 72; Mismatches 131; Indels 17; Gaps 4;

QY 12 ILAIGNCGGLIMRLYFNNGKRIWFSTFLETAGFPVIFIPLLFSYITRRSNNVGDST 71  
DB 7 VQVIGQSVAITILGRLYENNGSKWLAIVQVLGFPILLPYHLLSVKTHTTTQRDGLKTS 66

QY 72 PFLIKPRLLIAAIVIGILSGFDNLYAYGAYLVPSTAALIIASQLAFIAIFSPFMVKHK 131  
DB 67 ---LNRAVL-YIVLGLVGAACYLSYIGLLYLPVSTLSLICASQLAATFAPSYLNSQK 122

QY 132 FTPFTINAVLLTGAAVLGMHTETDKPVHETHKQYITGFLITVAAVMYAFILPLVELA 191

```
Db 123 LTPILNLSFLTLISSTLLAFNNEEDSKVTRGEYVKGFCVTVGASAGFGLLSLQOLA 182
QY 192 YKAKQKMTSYTLVLEFQLICLLASIVSVIGMFIAGDPKQALPKAREPKLGFALFVVA 251
Db 183 FRKVLKQKFSEVINMIYMSLVASCVSVGLFASSEWK-TLSEMEYKLGKVSVMNL 241
QY 252 VFSATIWQGFPGGALIGLFTSTSLVSGIMISVLLPITEVLAVIFYHEKFAEKGLSLALS 311
Db 242 VMTAVTWQVFSIGCTGLIFELSLFSLFSAISALGLPVVPIILAVIFHDKNKGLKVISMILA 301
QY 312 LMGFSVYEF-----GRIKSGEDKRRIQEESQE 339
Db 302 IMGFSVYVQQYLDETNLKKSNIEPTSPDRPEAGSSE 341

RESULT 5
T04922
hypothetical protein T9A21.50 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 11-Jun-1999
C:Accession: T04922
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15390
A:Accession: T04922
A:Molecule type: DNA
A:Residues: 1-1128 <BEV>
A:Cross-references: EMBL:AL021713
A:Experimental source: cultivar Columbia; BAC clone T9A21
C:Genetics:
A:Map position: 4
A:Introns: 76/2; 402/1; 526/3; 762/1
A:Note: T9A21.50

Query Match 31.6%; Score 557; DB 2; Length 1128;
Best Local Similarity 36.1%; Pred. No. 1.4e-34;
Matches 125; Conservative 76; Mismatches 131; Indels 14; Gaps 5;

QY 2 VKALVIINCIIAIGNCGPLIMRLYFNNGKRIWFSTFLETAGPVPVIFPLLFSTYITRR 61
Db 795 IYAFVIFQPLA-----TVGLRYENGKSKTYVVTLLQLIGFVPLILFRFESRIOP 848
QY 62 RS--NNVGDSSTFPLIKPRLIAAIVIGLSDNYLYAGIAYLPVSTAALIIASQLAF 119
Db 849 KSTDNTFNSQSPSTTLASVLCGLLVAYA----YLSAVGLYLPVSTFSLILASQLAF 904
QY 120 IAFSPFMVKHPTPTINAVLLTVGAALVGMHTETDKPVHETHKQYITGPLITVAAV 179
Db 905 TAFSFLNSQKPTPLIVNSLPLTVSSALLVNTDSENTNVSRYQYVIGFTCTIGASA 964
QY 180 MYAFILPLVELAYOKAKQMTSYTLVLEFQLICLLASIVSVIGMFIAGDPKQALPKARE 239
Db 965 GTGLVLSLLQLFRKVTKHTSVAVDLANYQSLVATCVVLIGLFASGBWR-TLPSEMRN 1023
QY 240 FKLGEALFYVAVFSAIIWQGFPLGALIGLFTSSLSVSGIMISVLLPITEVLAVIFYHEK 299
Db 1024 YKLGKVSYILTASAAIFQVYTVGCVGLIFESSVSFSNITAVGLPIVPVAVIVFHDK 1083
QY 300 FOAEKGLSLALSIMGVSVFYGEIKSGEDKRRIQEESQETBOSSL 345
Db 1084 MDASKIFSIILAIWGLPLSPVYQHYLD-EKKLTKCTQKPVVEETQTL 1128

RESULT 6
D96506
hypothetical protein T12C22.2 [imported] - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 02-Mar-2001 #sequence_revision 02-Mar-2001 #text_change 23-Mar-2001
C:Accession: D96506
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;
ansen, N.P.; Hughes, B.; Huizar, L.
```

```
Nature 408, 816-820, 2000
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.;
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.;
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.
A>Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.
A:Reference number: A86141; MUID:21016719; PMID:11130712
A:Accession: D96506
A>Status: preliminary
A:Molecule type: DNA
A:Residues: 1-379 <STO>
A:Cross-references: GB:AE005173; NID:98655985; PIDN:AAF78258.1; GSPDB:GN00141
C:Genetics:
A:Gene: T12C22.2
A:Map position: 1
C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 31.6%; Score 556.5; DB 2; Length 379;
Best Local Similarity 34.4%; Pred. No. 5.3e-35;
Matches 115; Conservative 89; Mismatches 121; Indels 9; Gaps 4;

QY 5 LVITNCIIAIGNCGPLIMRLYFNNGKRIWFSTFLETAGPVPVIFPLLFSTYITRRSN 64
Db 45 LVSVINIFFLIGGQASVLLGRFYDEGNSKMWATLVQTAAPFILIPLLL-----LPSSA 100
QY 65 NVGDSSTFPLIKPRLIAAIVIGLSDNYLYAGIAYLPVSTAALIIASQLAFAIPS 124
Db 101 SVESSESSCLXIVLI-YVLGVITAGDNMLYVGLLYLSASTYSLICATQAFNAVFS 159
QY 125 FFMVKHPTPTINAVLLTVGAALVGMHTETDKPVHETHKQYITGPLITVAAVMYAFI 184
Db 160 YFINAQKFTALILNSVLLSFAALIALNDADTPSGVSRKYIVGVCTLAASALYSLL 219
QY 185 LPLVELAYOKAKQMTSYTLVLEFQLICLLASIVSVIGMFIAGDPKQALPKAREPKLGE 244
Db 220 LSLMQFSPFKILKRETFSVVLEMQIYTSLVATCVSVIGLFASGEWR-TLHGEMEGYHKQ 278
QY 245 ALFYVAVFSAIIWQGFPLGALIGLFTSSLSVSGIMISVLLPITEVLAVIFYHEKFAEK 304
Db 279 ASVLTWLTWATVWQCVGVGLIFLVTLSFNSVISTLSLAVTPLAALVVPFRKMSGVK 338
QY 305 GLSLALSIMGVSVFYGEIKSGEDKRRIQEESQ 338
Db 339 IWAMLIATWGFASYVY---QNHIDDLKLVQARQ 369

RESULT 7
T04923
hypothetical protein T9A21.60 - Arabidopsis thaliana
C:Species: Arabidopsis thaliana (mouse-ear cress)
C>Date: 23-Apr-1999 #sequence_revision 23-Apr-1999 #text_change 21-Jul-2000
C:Accession: T04923
R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X.
submitted to the Protein Sequence Database, February 1999
A:Reference number: Z15390
A:Accession: T04923
A:Molecule type: DNA
A:Residues: 1-348 <BEV>
A:Cross-references: EMBL:AL021713
A:Experimental source: cultivar Columbia; BAC clone T9A21
C:Genetics:
A:Map position: 4
A:Introns: 24/2
A:Note: T9A21.60
C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 31.6%; Score 555.5; DB 2; Length 348;
Best Local Similarity 36.7%; Pred. No. 5.8e-35;
Matches 120; Conservative 72; Mismatches 120; Indels 15; Gaps 6;

QY 26 LYFNNGKRIWFSTFLETAGPVPVIFPLLFSTYITRRSNVNGDSTSFLLIKPR-LLIAAV 84
```







Mon Sep 27 08:51:17 2004

ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: H96516  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-392 <STO>  
A:Cross-references: GB:AE005173; NID:G5668799; PIDN:AAD46025.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F16N3.10  
A:Map position: 1

Query Match 15.9%; Score 279.5; DB 2; Length 392;  
Best Local Similarity 25.1%; Pred. No. 6e-14;  
Matches 90; Conservative 74; Mismatches 153; Indels 41; Gaps 13;

QY 5 LVINCIILAIAGCGGLIMRLYFNNGKRI-----WFSTFLETAGFP-VIFPL 53  
DB 48 IFVCSLWVA-GRVLSTLLNFYFIQTGRDVCDDPKQFKGTWQWQNAAPPTAFLL 106  
QY 54 LFSYITRRSNNVGDSTFFLIKPRLLIAAVVILSGFDNYLYAYG-----IAYLPVSTA 109  
DB 107 LWRSSFSTHSTSSSSSF---GKFLLYISLGLVLAAYSLYAGRTGTHCVFLL----- 157  
QY 110 ALIIASOLAFIAFFFWKHKFTFTINAVVLLTVGAALVGMHMTETDK-PVHETHKQ 166  
DB 158 -WIFTSQIFTSIFTIINKQFNWILSLMVL--SGAATGLGITSSGGAVIPCENEGSK 214  
QY 167 YITGLITVAAMVAFILPLVELAYQAKQMS--YTLVLEFQILCLILASIVSVMGF 224  
DB 215 MNGAWCAFFGTVAFLSLCIMQLGFKQVPTQSRVSAVILMQTNASMIATLCLVGLF 274  
QY 225 IAGDFKQALPKAREFKLGEALFYVAVFSAIHWQFFLGAIGLIFSTSSLSVSGIMISVL 284  
DB 275 VSGEFKD-IKEDFETFKGKPLVLSLGLSLAWQVMSLGLVGLVGLVGLVGLVGLV 333  
QY 285 LPITEVLAVIFVH-----EKFOAEKGLSLALSLWGFVSFYFYGEIKSGEDKRRIQEES 337  
DB 334 TELVNILLVLAFRFTDADVKFFKEG--ALVAGILGFASVYSLYKSTKKEIASOSQT 389

RESULT 14  
E96516  
F16N3.13 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: E96516  
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: E96516  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-389 <STO>  
A:Cross-references: GB:AE005173; NID:G5668802; PIDN:AAD46028.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F16N3.13  
A:Map position: 1

Query Match 15.8%; Score 277.5; DB 2; Length 389;  
Best Local Similarity 24.9%; Pred. No. 8.4e-14;  
Matches 90; Conservative 68; Mismatches 151; Indels 53; Gaps 12;

QY 10 CIILAIAGCGG-----PLIMRLYFNNGG-----KRIWFSTFLETAGFPVIF 50  
DB 47 CIFV-----CSGFVVTGRVLSLTLNLYFIQTGRDACDDPKQFKGTWQWQNAAPPTAF 102  
QY 51 IPIPLFYITRRSNNVGDSTFFLIKPRLLIAAVVILSGFDNYLYAYGIAVLPVSTAA 110  
DB 103 FLILLWRSLSFSTHGETQSSSF---GKFLLYISLGLV--FSAYSLYAGRTGTHCVF 156  
QY 111 LIITASOLAFIAFFFWKHKFTFTINAVVLLTVGAALVGMHMTETDKPVHETHKQYITG 170  
DB 157 WITTLQIFTSIFTAINKHKNRWILSVLSGVATGI-----TSSDDAYVPCSE---G 209  
QY 171 FLITVAA-----AVMYAFILPLVELAYQAK--KQMSYTLVLEFQILCLILASIVSVIG 222  
DB 210 WKMSYGAWCSFFGTVAFLSLCIMQLGFKQVPTQSRVSAVILMQTNASMIATLCLV 269  
QY 223 MFTAGDFKQALPKAREFKLGEALFYVAVFSAIHWQFFLGAIGLIFSTSSLSVSGIMIS 282  
DB 270 LFSVSGEFKD-IKEDFETFKGKPLVLSLGLSLAWQVMSLGLVGLVGLVGLVGLV 328  
QY 283 VLLPITEVLAVI---FYHEKFOAEKGLSLALSLWGFVSFYFYGEIKSGEDKRRIQEES 339  
DB 329 SATPVANIFVLAFRFMDDDIGWFKGALLGILGFASVYSLYKAIK-----KQEI 383  
QY 340 TE 341  
DB 384 TE 385

RESULT 15  
C96613  
hypothetical protein F13D13.4 [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: C96613  
R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziani,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: C96613  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-398 <STO>  
A:Cross-references: GB:AE005173; NID:g1128382; PIDN:AAG31188.1; GSPDB:GN00141  
C:Genetics:  
A:Gene: F13D13.4  
A:Map position: 1

Query Match 14.9%; Score 262.5; DB 2; Length 398;  
Best Local Similarity 26.5%; Pred. No. 1.2e-12;  
Matches 85; Conservative 67; Mismatches 142; Indels 27; Gaps 10;

QY 33 KRIWFSTFLETAGFPVIFIPLLFSYITRRSNNVGDSTFFLIKPRLLIAAVVILSGIF 92  
DB 85 KGTWQALIQNAAPFIL-IPLFFIF-PRKQHLNNTSFLSL--RLFFLYLSLGLVAA 140  
QY 93 DNYLVAYGIAVLPVSTAAALIIASOLAFIAFFFWKHKFTFTINAVVLLTVGAALVGM 152  
DB 141 HSKLFALGKLVSNYGIFSLISTTQIFTAIVLTAINRFRKFRWILIS-ILITIVYVIGT 199  
QY 153 HTETDKPVHETHKQYITGLITVAAMVAFILPLVELAYQAK-QTMSY-----TLVL 205



```

Db      200 PDFGGQPHDGEERGYNIOAWLAFSATIAFSLSLCLIOLGPEKLOVTKRYGNEKVFRRMVL 259
Qy      206 EFQILCLLASIVSVMFIAGDFKQALPKAREFKLGEALFYVAVFSAIIWQGFELGA 265
      | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
Db      260 EMQICVAFVAVVCLVGLFASDEYKE-LKGDSCRFFKGGETYVVLVSLVGLALSQVWAVGM 318
      | | : | | : | | : | | : | | : | | : | | : | | : | | : | | : |
Qy      266 IGLIFSTSSIVSGIMISVLLPITEVLAV-----IFYHEKFOAEKGLSLALSLSLWGFVS 317
      | | | | | | | | | | : | | : | | : | | : | | : | | | | | | |
Db      319 IGLVHVYVSLFGDVVHMCASFFVALFVVLAFDFMDDVFSWPRIGALIGTVLALG-----S 373
      | | | | | | | | | | : | | : | | : | | : | | : | | | | | | |
Qy      318 YFYGEIKSGEDK-RRIOQBEES 337
      | | | | | | | | | | : | | : | | : | | : | | : | | | | | | |
Db      374 YFYTLHKRNKKKMAELNQSEN 394
      | | | | | | | | | | : | | : | | : | | : | | : | | | | | | |

```

Search completed: September 23, 2004, 23:51:18  
 Job time : 26.8588 secs

**THIS PAGE BLANK**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:42:01 ; Search time 25.1412 Seconds  
(without alignments)  
1362.072 Million cell updates/sec

Title: US-09-913-767-8  
Perfect score: 1798  
Sequence: 1 MNGLIIIIICITIGTCG.....DKQPPEPTELPLPVSVDYA 356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 283366 seqs, 96191526 residues

Total number of hits satisfying chosen parameters: 283366

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : PIR 78.\*  
1: pir1.\*  
2: pir2.\*  
3: pir3.\*  
4: pir4.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1792	99.7	356	2 D86408	purine permease [i
2	1160	64.5	356	2 B84749	hypothetical prote
3	1134.5	63.1	351	2 C86408	F3H9.12 protein -
4	550	30.6	379	2 D96506	hypothetical prote
5	545.5	30.3	344	2 T04924	hypothetical prote
6	543	30.2	348	2 T04923	hypothetical prote
7	533.5	29.7	1128	2 T04922	hypothetical prote
8	531.5	29.6	315	2 A84634	hypothetical prote
9	494.5	27.5	358	2 T04921	hypothetical prote
10	473.5	26.3	422	2 C85087	hypothetical prote
11	327	18.2	393	2 G86330	hypothetical prote
12	313	17.4	389	2 E96516	F16N3.13 [importe
13	287	16.0	390	2 D96613	unknown protein T1
14	282.5	15.7	392	2 H96516	F16N3.10 [importe
15	271	15.1	398	2 C96613	hypothetical prote
16	258	14.3	383	2 H86232	hypothetical prote
17	141	7.8	470	2 C69725	phosphotransferase
18	140	7.8	357	2 T50597	related to UDP N-A
19	138	7.7	368	2 T10557	hypothetical prote
20	134.5	7.5	343	2 C69212	conserved hypotet
21	132	7.3	311	2 F83698	hypothetical prote
22	128	7.1	370	2 A96507	hypothetical prote
23	126	7.0	298	2 A83892	hypothetical prote
24	125.5	7.0	348	2 AE1193	conserved hypotet
25	125	7.0	288	2 F72219	conserved hypotet
26	125	7.0	292	2 D71181	conserved hypotet
27	124	6.9	287	2 G69766	conserved hypotet
28	122	6.8	297	2 A83049	hypothetical prote
29	121	6.7	629	2 S29685	retroviral recepto

30	120.5	6.7	284	2 E71174	hypothetical prote
31	120.5	6.7	325	2 D75555	conserved hypotet
32	120.5	6.7	370	2 S69718	hypothetical prote
33	120.5	6.7	389	1 C64923	chloramphenicol re
34	120	6.7	287	2 E75159	hypothetical prote
35	120	6.7	310	2 C97242	probable permease
36	120	6.7	357	2 H84792	nodulin-like prote
37	119.5	6.6	345	2 AC1252	conserved hypotet
38	119	6.6	286	2 JS0384	hypothetical 30.5K
39	118	6.6	479	2 JQ0781	sucrose uptake pro
40	117.5	6.5	345	2 AI1614	conserved hypotet
41	116.5	6.5	347	2 G96741	unknown protein F1
42	116.5	6.5	389	2 F90924	probable transport
43	116.5	6.5	389	2 B85773	probable transport
44	116	6.5	412	2 AB2471	hypothetical prote
45	115.5	6.4	362	2 T05133	hypothetical prote

ALIGNMENTS

RESULT 1

D86408  
purine permease [imported] - Arabidopsis thaliana  
C:Species: Arabidopsis thaliana (mouse-ear cress)  
C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C:Accession: D86408  
R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon, I.  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A:Reference number: A86141; MUID:21016719; PMID:11130712  
A:Accession: D86408  
A:Status: preliminary  
A:Molecule type: DNA  
A:Residues: 1-356 <STO>  
A:Cross-references: GB:AE005172; NID:g9795614; PIDN:AAF98432.1; GSPDB:GN00141  
C:Genetics:  
A:Map position: 1  
C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 99.7% ; Score 1792; DB 2; Length 356;  
Best Local Similarity 99.7% ; Pred. No. 1.3e-135;  
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

Qy	1	MKNGLIICIITIGTCG	PLTTRLYFTNGGKRIWFM	SFLSTAGPFIILPLLVFLSR	60
Db	1	MKNGLIICIITIGTCG	PLTTRLYFTNGGKRIWFM	SFLSTAGPFIILPLLVFLSR	60
Qy	61	RRGNRPNNAENKRTK	LFMETPLFIASVIGLLTGLD	NYLSYGLAYLPVSTSSLIIG	120
Db	61	RRGNRPNNAENKRTK	LFMETPLFIASVIGLLTGLD	NYLSYGLAYLPVSTSSLIIG	120
Qy	121	TQLAFNALFAFLV	KQFTPFPSINAVVLLTVG	IGILALHSDGDKPAKESK	YVVGFLMT 180
Db	121	TQLAFNALFAFLV	KQFTPFPSINAVVLLTVG	IGILALHSDGDKPAKESK	YVVGFLMT 180
Qy	181	VVAALYAFILPLVEL	TYKARQEIFPLVLEIQM	VCLAAFFCVIGMFI	VGDFKVIAR 240
Db	181	VVAALYAFILPLVEL	TYKARQEIFPLVLEIQM	VCLAAFFCVIGMFI	VGDFKVIAR 240
Qy	241	EAREFKIGSVFY	YALIVITGIWQGFLLGA	IGIVFCASSLASGLSV	LISVLLPTEYFAVV 300
Db	241	EAREFKIGSVFY	YALIVITGIWQGFLLGA	IGIVFCASSLASGLSV	LISVLLPTEYFAVV 300
Qy	301	CFREKFAEKGV	SLLSLMGVSFYFYGE	FKSGKKVVDKFPPE	TELPILPVSVDYA 356

Db 301 CFREKFOAEKGVSLLSLWGFVSFYGFSGKGVKVPQPPETELPILPVSDIYA 356

RESULT 2

B84749

hypothetical protein At2g33750 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: B84749

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Unayam, L.; Tallon, L.; Euse, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.B.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: B84749

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-356 <STO>

A:Cross-references: GB:AE002093; NID:g1707019; PIDN:AAC69140.1; GSPDB:GN00139

C:Genetics:

A:Gene: At2g33750

A:Map position: 2

C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 64.5%; Score 1160; DB 2; Length 356;

Best Local Similarity 62.9%; Pred. No. 3.5e-85;

Matches 231; Conservative 45; Mismatches 69; Indels 22; Gaps 4;

QY 1 MKNGLIINCIIITGTCGGPLTRLYFTNGGKRWFMFSLSTAGFPILPLVSPLSR 60

DB 1 MKTVLIINCIIAIGNCGGLMRLYFQNGERWFFSLQTVCGPLIFFPLLSFLRR 60

QY 61 RGRNRPNNNAENKRTKFLMETPLFIASIVIGLITGLDNLVSYGLAYLPVSTSLIIG 120

DB 61 RR-----CLEQSEITPFLMPKPFIAIVGLLVGFDNLYSYGLAYIPVSTASLIIS 114

QY 121 TQLAFNALFAFLVVKQKFTPFSSINAVLLTVGIGIILALHSDGDKPAKESKKEVYVGFMT 180

DB 115 AQLGTFALFAFMVKQKFTPTINATVLLTGGAVLALNSDSKLANETHKEVYVGFMT 174

QY 181 VVAALLYATILPVELTYKQKQETFPFLVLEIOMVCLAAATFFCVIGMIFVDPKVIAR 240

DB 175 LGAALLYGLPLVLSYKSGGRIYITLALFQWVLCFAATCVLGVMLAAGDKFKVHA 234

QY 237 -----VIAREAREFKIGSVFYVALIVITIGIIGQFFLGAIGIVFCASSLASGLVLSV 289

DB 235 LFTFKNRVITAGEARDFKLGESL-YVYVIVFTAIWQAFVGAIGLIFCASSLSVGMVSA 293

QY 290 LLPVTEVFAVCPREKFOAEKGVSLLSLWGFVSFYGFSGKGVKVPQPPETELPIL 349

DB 294 LLPVTIVLAVICQERFQAGKVALALSLSWGSVSFYGVQVKS-----EETKAQDTQLSOL 349

QY 350 PVSDIYA 356

DB 350 PVTDIYA 356

RESULT 3

F3H9.12 protein - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 27-Nov-2001

C:Accession: C86408

R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, R.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96506

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-379 <STO>

A:Cross-references: GB:AE005173; NID:g9655985; PIDN:AAF78258.1; GSPDB:GN00141

C:Genetics:

A:Gene: TL2C22.2

A:Map position: 1

C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 30.6%; Score 550; DB 2; Length 379;

Best Local Similarity 34.4%; Pred. No. 1.7e-36;

Matches 119; Conservative 74; Mismatches 133; Indels 20; Gaps 4;

QY 5 LIIINCIIITGTCGGPLTRLYFTNGGKRWFMFSLSTAGFPILPLVSPLSR 64

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C86408

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-351 <STO>

A:Cross-references: GB:AE005172; NID:g9795615; PIDN:AAF98433.1; GSPDB:GN00141

C:Genetics:

A:Map position: 1

C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 63.1%; Score 1134.5; DB 2; Length 351;

Best Local Similarity 59.9%; Pred. No. 3.7e-83;

Matches 214; Conservative 66; Mismatches 66; Indels 11; Gaps 4;

QY 1 MKNGLIINCIIITGTCGGPLTRLYFTNGGKRWFMFSLSTAGFPILPLVSPLSR 60

DB 1 MVRKALVIINCIIAIGNCGGLMRLYFNNNGKRWFFSTFLETAGFPFVIFIPLLFSYIT 60

QY 61 RGRNRPNNNAENKRTKFLMETPLFIASIVIGLITGLDNLVSYGLAYLPVSTSLIIG 120

DB 61 RRSNNVGD-----TSFFLIKPRLLIAAVIVGILSGFDNLYAYGLAYLPVSTAALIIA 114

QY 121 TQLAFNALFAFLVVKQKFTPFSSINAVLLTVGIGIILALHSDGDKPAKESKKEVYVGFMT 180

DB 115 SQAFAITAFSFFWVGHKFTPTINAVLLTVGAALVGMHTETDKPVHETHKQVITGFLIT 174

QY 181 VVAALLYATILPVELTYKQKQETFPFLVLEIOMVCLAAATFFCVIGMIFVDPKVIAR 240

DB 175 VAAAVMYAFILPLVELAYQAKQMTSYTLVLEFQLILCLLASIVSVIGMFIAGDFKALPK 234

QY 241 EAREFKIGSVFYVALIVITIGIIGQFFLGAIGIVFCASSLASGLVLSVLLPVTVEFAV 300

DB 235 EAREFKLGEALF-YVAVFSAIIVQGFGLGALGIFSTSSILVSGIMISVLLPITEVLAVI 293

QY 301 CFREKFOAEKGVSLLSLWGFVSFYGFSGKGVKVPQPPETELPIL--PVSD 353

DB 294 FYHEKFOAEKGLSLSLWGFVSFYGFSGKGVKVPQPPETELPIL--PVSD 350

RESULT 4

D96506

hypothetical protein TL2C22.2 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 23-Mar-2001

C:Accession: D96506

R:Theologias, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso, C.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, R.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Creasy, T.H.; Dewar, K.; ansen, N.F.; Hughes, B.; Huizar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C. C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali, Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: D96506

A:Status: preliminary

A:Molecule type: DNA

A:Residues: 1-379 <STO>

A:Cross-references: GB:AE005173; NID:g9655985; PIDN:AAF78258.1; GSPDB:GN00141

C:Genetics:

A:Gene: TL2C22.2

A:Map position: 1

C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 30.6%; Score 550; DB 2; Length 379;

Best Local Similarity 34.4%; Pred. No. 1.7e-36;

Matches 119; Conservative 74; Mismatches 133; Indels 20; Gaps 4;

QY 5 LIIINCIIITGTCGGPLTRLYFTNGGKRWFMFSLSTAGFPILPLVSPLSR 64

Db 45 LVSNIFFLIGGQAASVLLGRFYDEGGNKKMATLVQTAAPFLYIPLLL-----LPSS 99

QY 65 RNFNNAENKRTKLFMETPLFIASIVIGLITGLDNLVSYGLAYLPVSTSLIIGTOLA 124

Db 100 ASVESSESSCLKY-----IVLIYVLLGVIIAGDNLVSYGLLISASTYSLICATOLA 153

QY 125 FNALPAPLLVKKQETPPSINAVLLTVIGIGILAHSDGDKPAKESKKEYVVGFLMTVAA 184

Db 154 FNAVFSYFNAQKFTALILNSVLLSFAALIALNDADTFPSGVRSKYIVGVFCTLAAS 213

QY 185 LLYAFILPLVELTVKKARQBITPLVLEIOWMCLAAATFCVIGMFIVGDFKVIAREARE 244

Db 214 ALYSLLLSLMQFSEKILKETSTVLEMQIYSLVATCVSVIGLFGASGWRKTLHMEG 273

QY 245 FKIGGSVYVALIVITGIIWQGFPLGAIGVFCASSLASGVLISVLLPVTVEFAVVCFRE 304

Db 274 YH-KGQASYVLTAVTAVTQVCSGVVGLIFLWTSFNSVISTLSLAVTFLAALVVRD 332

QY 305 KFOAKGVSLLSLWGFVSYPYGFESGKKV-----VDKQPPP 342

Db 333 KMSGVKIMAMLIAIWGPASYVYQNHDDLVKVRQAQQAQGRVEPP 378

RESULT 5

T04924

hypothetical protein T9A21.70 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 20-Jun-2000

C;Accession: T04924

R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X

submitted to the Protein Sequence Database, February 1999

A;Reference number: Z15390

A;Molecule type: DNA

A;Residues: 1-344 <BEV>

A;Cross-references: EMBL:AL021713

A;Experimental source: cultivar Columbia; BAC clone T9A21

C;Genetics:

A;Map position: 4

A;Introns: 11/1

A;Note: T9A21.70

C;Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 30.3%; Score 545.5; DB 2; Length 344;

Best Local Similarity 35.1%; Pred. No. 3.6e-36;

Matches 120; Conservative 69; Mismatches 142; Indels 11; Gaps 2;

QY 12 ILTIGCGPLLRLTYFTNGGKRIWFMSTAGFPILIPLLVSLFSLRRGRNPNNAE 71

Db 7 VQVIGOSVATILGRLYENGNSKWLATVQVLPVLLPYHLLSVKTHTTTQORDGKLS 66

QY 72 NKRKTKLFMETPLFIASIVIGLITGLDNLVSYGLAYLPVSTSLIIGTOLAPNALPAP 131

Db 67 LNRALVY-----IVIGLVGAACYLSIGLLYLPVSTSLICASQAFAPFSY 116

QY 132 LLVKQKFTPPSINAVLLTVIGIGILAHSDGDKPAKESKKEYVVGFLMTVAAALLYAFIL 191

Db 117 LLNSQKLTPIILNSFLITISLTLLAFNNEESDKVTKGEYKGVFCTVGSAGFGLLL 176

QY 192 PLVELYKKARQBITPLVLEIOWMCLAAATFCVIGMFIVGDFKVIAREAREFKIGGSV 251

Db 177 SLQOLAFKVLKQKQTFSEVINMIYMSLVASCVSVGLFASSEWKTLSSEMYNKL-GKV 235

QY 252 FYVALIVITGIIWQGFPLGAIGVFCASSLASGVLISVLLPVTVEFAVVCFREKFOAKG 311

Db 236 SYVMNLVAVTAVTQVFSIGCTGTLFELSLFNSALGLPVPVFLAVIIFHDKMGLKV 295

QY 312 VSLLSLWGFVSYPYGFESGKKVWDKQPPPETELPILPVSD 353

Db 296 ISWILAIWGFVSIVYQQYLDLTKKSNELPTTESDPPEAE 337

RESULT 6

T04923

hypothetical protein T9A21.60 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-Jul-2000

C;Accession: T04923

R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X

submitted to the Protein Sequence Database, February 1999

A;Reference number: Z15390

A;Accession: T04923

A;Molecule type: DNA

A;Residues: 1-348 &lt;BEV&gt;

A;Cross-references: EMBL:AL021713

A;Experimental source: cultivar Columbia; BAC clone T9A21

C;Genetics:

A;Map position: 4

A;Introns: 24/2

A;Note: T9A21.60

C;Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 30.2%; Score 543; DB 2; Length 348;

Best Local Similarity 34.4%; Pred. No. 5.7e-36;

Matches 116; Conservative 69; Mismatches 130; Indels 22; Gaps 4;

QY 17 TCGPPLLRLTYFTNGGKRIWFMSTAGFPILIPLLVSLFSLRRGRNPNNAENKRT 76

Db 19 TNGG---SGVYDNGGNSKWLATVQVLPVLLPYIILSPKTHATTDGKRTSPNRV 75

QY 77 KLFMETPLFIASIVIGLITGLDNLVSYGLAYLPVSTSLIIGTOLAFNALPFLVVK 136

Db 76 LVV-----VLGLVGADCYLSIGLLYLPVSTSLICASQAFNAPFSYLSNQ 125

QY 137 KTFPFSINAVLLTVIGIGILAHSDGDKPAKESKKEYVVGFLMTVAAALLYAFILPVEL 196

Db 126 KLTPIILNSFLITISLTLLAFNNEETDSTKVTKGEYKGVFCTVVASAGYGLVLSQOL 185

QY 197 TYKARQBITPLVLEIOWMCLAAATFCVIGMFIVGDFKVIAREAREFKIGGSVFFYAL 256

Db 186 AFLKVLKQNFSEVMDMIYVSLVASCVSVGLFASSEWKTLSSEMDNYK-HGKVSYIMN 244

QY 257 IVITGIIWQGFPLGAIGVFCASSLASGVLISVLLPVTVEFAVVCFREKFOAKGVSLLL 316

Db 245 LWMTAVTQVFSIGGTGLFELSLFNSALSVGLPVPVFLAVIIFHDKMGLKVSIMIL 304

QY 317 SLWGFVSYPYGFESGKKV-----VDKQPPPETE 345

Db 305 AIWGFYSYVYQQYLDLTKKSNELPTTESDPPEAE 341

RESULT 7

T04922

hypothetical protein T9A21.50 - Arabidopsis thaliana

C;Species: Arabidopsis thaliana (mouse-ear cress)

C;Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 11-Jun-1999

C;Accession: T04922

R;Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X

submitted to the Protein Sequence Database, February 1999

A;Reference number: Z15390

A;Accession: T04922

A;Molecule type: DNA

A;Residues: 1-1128 &lt;BEV&gt;

A;Cross-references: EMBL:AL021713

A;Experimental source: cultivar Columbia; BAC clone T9A21

C;Genetics:

A;Map position: 4

A;Introns: 76/2; 402/1; 526/3; 762/1

A;Note: T9A21.50

Query Match 29.7%; Score 533.5; DB 2; Length 1128;

Best Local Similarity 34.8%; Pred. No. 1e-34;

Matches 114; Conservative 72; Mismatches 133; Indels 9; Gaps 4;

QY 26 LYFTNGGKRIWFMSTAGFPILIPLLVSLFSLRRGRNPNNAENKRTKLFMLMETPL 85

Db 77 VYENGKSTWMTGLVQLIGFPVL---FLRFFSQT-KNPKPTADPRK-----PSSFTIL 128

Qy 86 FIASVIGLGLDNLVSYGLAYLPVSTSSLIIGTQLAFNALFAFLLVKQKTPFFSINA 145

Db 129 GSVYIVTGLLVANSYMSVGLVLPVSTFSLILASQLAFTAFSYFLNSQKTPFFIVNS 188

Qy 146 VLLVTVGIIALHSDGDKPAKESKKEVVGFLMTVVAALYAFILPLVELTYKARQEI 205

Db 189 LFLFTISSALLVWNTDSENTAKVRKVYVIGICTIGASAGIGLLSLVQLILKVLKQ 248

Qy 206 TPPLVLEIQMWCLAAATFCVIGMFVDFVIAREAREFKIGSVFYVALIVITGIWQ 265

Db 249 TFSVTVDLVAYQSLVASCVVLLGLFASGEWKLITSEMENYKL-GKVPVWTLASIASWQ 307

Qy 266 GFELGAIGIVFCASSLASGLISVLLPVTVEFAVVCFEKFOAEKGVSLLSLGMFVSF 325

Db 308 VYIGVGLIPESSSVFSNSITAVGLPVPVAVIVFDHKNASKIFSIIILAINGFISFV 367

Qy 326 YGEFKSGKVKVDKQPPPTLPILPVSQ 353

Db 368 YQHYLDEKKLKTSHTSFVGDPHLLPAEE 395

RESULT 8

A84634

hypothetical protein At2g24220 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 02-Feb-2001 #sequence\_revision 02-Feb-2001 #text\_change 16-Feb-2001

C:Accession: A84634

R:Lin, X.; Kaul, S.; Rounsley, S.D.; Shea, T.P.; Benito, M.I.; Town, C.D.; Fujii, C.Y.; M.; Koo, H.; Moffat, K.S.; Cronin, L.A.; Shen, M.; Vanaken, S.E.; Umayam, L.; Tallon, L.; euss, D.; Nierman, W.C.; White, O.; Eisen, J.A.; Salzberg, S.L.; Fraser, C.M.; Venter, J. Nature 402, 761-768, 1999

A:Title: Sequence and analysis of chromosome 2 of the plant Arabidopsis thaliana.

A:Reference number: A84420; MUID:20083487; PMID:10617197

A:Accession: A84634

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-315 <STO>

A:Cross-references: GB:AE002093; NID:g4115382; PIDN:AAD03383.1; GSPDB:GN00139

A:Gene: At2g24220

A:Map position: 2

C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 29.6%; Score 531.5; DB 2; Length 315;

Best Local Similarity 36.8%; Pred. No. 4.3e-35;

Matches 114; Conservative 67; Mismatches 114; Indels 15; Gaps 2;

Qy 22 LLRLYFTNGGKRIPWMSFLSTAGFPFIILIPVLSFLRRGRNPNNAENKRTKILFLM 81

Db 10 LLSRLYFNGGKSKWIIISWVAVAGNPITCLILLTYIF-----QKIKPTPL 55

Qy 82 ETPLFTASVIGLGLDNLVSYGLAYLPVSTSSLIIGTQLAFNALFAFLLVKQKTPFP 141

Db 56 NTKLVLSYVVLGFLSADNLNMAVAYAYLPASTSSLASLAFSALFGYLVKKNPLAS 115

Qy 142 SINAVLLTVGIGILALHSDGDKPAKESKKEVYVGLFMTVVAALYAFILPLVELTYKA 201

Db 116 VINSIVITGMAIILDSSSDRYISNSQYFAGFFWIDMSALHGLIFALSSELLFVKL 175

Qy 202 RQBITPLVLEIQMWCLAAATFCVIGMFVDFVIAREAREFKIGSVFYVALIVITG 261

Db 176 LGRSFHFVALEQQVMSLTAFATTTIGMVVNSDFQGMSEAKSF-KGGSLEYQVLWNSA 234

Qy 262 IIQGFFLGAIGIVFCASSLASGLISVLLPVTVEFAVVCFEKFOAEKGVSLLSLWGF 321

Db 235 VTQLGVLGATVLFVASTWAGVLANVRVPIITSVAAILMHPDMSGFKILSLVLTFFWGF 294

Qy 322 VSYFYGEFKS 331

Db 295 SSVYIGSSSS 304

RESULT 9

T04921

hypothetical protein T9A21.30 - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 23-Apr-1999 #sequence\_revision 23-Apr-1999 #text\_change 21-Jul-2000

C:Accession: T04921

R:Bevan, M.; Murphy, G.; Ridley, P.; Hudson, S.; Bancroft, I.; Mewes, H.W.; Mayer, K.F.X. submitted to the Protein Sequence Database, February 1999

A:Reference number: Z15390

A:Accession: T04921

A:Molecule type: DNA

A:Residues: 1-358 <BEV>

A:Cross-references: EMBL:AL021713

A:Experimental source: cultivar Columbia; BAC clone T9A21

C:Genetics:

A:Map position: 4

A:Introns: 60/3

A:Note: T9A21.30

C:Superfamily: Arabidopsis thaliana hypothetical protein T9A21.70

Query Match 27.5%; Score 494.5; DB 2; Length 358;

Best Local Similarity 34.6%; Pred. No. 4.4e-32;

Matches 118; Conservative 71; Mismatches 119; Indels 33; Gaps 8;

Qy 12 ILTIGTCGGPLLRLYFTNGGKRIPWMSFLSTAGFPFIILIPVLSFLRRGRNPNNAE 71

Db 29 LLLAGETIATLLGLRLYKGGKSTW----LET----LVQLP-----EPSTK 67

Qy 72 --NKRRTKLFIMETPLFIASIVIGLITGLDNLVSYGLAYLPVSTSSLIIGTQLAFNALF 129

Db 68 TIRKTTSSFL---TSLVYIGLGLVAGHCILYFGLLYLPVSTFSLISASQLAFNAVF 124

Qy 130 AFLVVKQKTPFNSINAVLLTVGIGILALHSDGDKPAKES----KKEVYVGLMTVVAAL 185

Db 125 SYFLNSQKITPFIILNSLVLLTISTLLVIOHEPESPSTSKSAKSKYVIGYICAVGSSA 184

Qy 186 LYAFILPLVELTYKARQEIITFPLVLEIQMWCLAAATFCVIGMFVDFVIAREAREF 245

Db 185 GYSLVLSLTADYAFKILKKYTRAILDMATVPMSVATCVVVGFLFGSGGKKLSTEMEER 244

Qy 246 KIGGSVFYALIVI-TGIWQGFPLGAIGIVFCASSLASGLISVLLPVTVEFAVVCFRE 304

Db 245 QLGS--SYILINIGSTISWQALIGSVGLIIEVSLFSNVISTLCPLVPVVPVLAUVFFRD 302

Qy 305 KFOAEKGVSLLSLWGFVSFYGEFKSGKVKVDKQPPPETE 345

Db 303 EMSGIKLVAMPATWGFVSYGQHYVNDRKPEEDQELPQSK 343

RESULT 10

C85087

hypothetical protein AT4g08700 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C>Date: 16-Feb-2001 #sequence\_revision 16-Feb-2001 #text\_change 16-Feb-2001

C:Accession: C85087

R:anonymous, The European Union Arabidopsis Genome Sequencing Consortium, The Cold Spring Nature 402, 769-777, 1999

A:Title: Sequence and analysis of chromosome 4 of the plant Arabidopsis thaliana.

A:Reference number: A85001; MUID:20083486; PMID:10617198

A:Accession: C85087

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-432 <STO>

A:Cross-references: GB:NC\_001268; NID:g7267512; PIDN:CAB77995.1; GSPDB:GN00140

C:Genetics:

A:Gene: AT4g08700

A:Map position: 4

Query Match 26.3%; Score 473.5; DB 2; Length 432;

Best Local Similarity 32.6%; Pred. No. 2.5e-30;

Matches 108; Conservative 73; Mismatches 135; Indels 15; Gaps 4;

QY 5 LIINCILLITGCGPPLRLRYLTNGGKRWMFSLTAGFPILPLLVSLFSLRRGN 64  
Db 31 LVFISIFLISAQIAIALLGRFYNEGSKWISTLVQTGCPILYPL--CFLPA--- 84  
QY 65 RNPNAENKRTKLFMETPLFIASIVGLTGLDNLVSLYGLAYLPVSTSLIIGTOLA 124  
Db 85 SHSSSSCSFKTLWYI-----LSLGFAGLDNLVSLYGLAYLPVSTSLIIGTOLA 136  
QY 125 FNALFAFLVKKQFTPFPSINAVLLTVIGIGILALHSDGDKPAKESKKEYYVVGFLMTVAA 184  
Db 137 FNGVFSYINSQKITCLIFSLVSLVSAVLSDDDNSPGSDKWSYLGICLCTVFAS 196  
QY 185 LLYAFILPLVELYTKARQBITPLVLEIQNMVCLAAATFFCVGMFTVIGDKVIAREARE 244  
Db 197 LIYSLQLSLMQFSENVKSETFSVLEMLQYITSLVASCVAIVGLFASGEWMLLSVEMEE 256  
QY 245 FKIGGSVYVALIVITGIIWQGFPLGAIGIVFCASSLASGLVSLVLLPVTVEFVAVCFRE 304  
Db 257 FH-EGQVYVLTIVGTVAVSQVGLSVGAVLIFLVSSLSNLIGTSLIVTPLAAIAVFHD 315  
QY 305 KFOAEKGVSLLSLWGFVSYFYGFKSGKKV 335  
Db 316 KLTEVKNVAMLIAPMGFGFYIYQNYLDDLKV 346  
RESULT 11  
G86330  
Query Match 17.4%; Score 313; DB 2; Length 389;  
Best Local Similarity 28.0%; Pred. No. 1.4e-17;  
Matches 99; Conservative 60; Mismatches 150; Indels 42; Gaps 10;  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Dec-2001  
C;Accession: G86330  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: G86330  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-389 <STO>  
A;Cross-references: GB:AE005173; NID:g5668802; PIDN:AAD46028.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: Fl6N3.13  
A;Map position: 1

QY 6 IIINCILLITGCGPPLRLRYL--TNGGK-----RIWFMFSLTAGFPILPLLV 55  
Db 49 ITISIIIFIQSIKLENNFYDKTNSEYNNRQNDGWVTSQSLTGTGFPPLLLPFLI 108  
QY 56 SFLSRGRGNPNNAENKRTKLFMETPLFIASIVIGLTLGLDNLVSLYGLAYLPVSTS 115  
Db 109 FITKNR-----NHHQPPITSDSIHLKSLAVIYICIGIIMSQQRLAAMGKLEIPFGVF 163  
QY 116 SLIIGTQLAFNALPAFLVKKQFTPFPSINAVLLTVIGIGILALHSDGDKPAKESKKEYYV 175  
Db 164 TLIYTAQLFTPTIAPFINKIKENRWVVISVILAIL--TGALTUSSSPGSPDRAENYAR 222  
QY 176 GLMTVVAALYAFILPLVELTY-----KKARQBITFPLVLEIQNMVCLAAATFFCVI 227  
Db 223 GSWAALFAGICFALLCNQNVDSYIFKTESTNQKPSFASVFEVIFPSLVAITISV 282  
QY 228 GMIVGDKVIAREAREFKTG-GSVFYVALIVITGIIWQGFILGAIGIVFCASSLASGLV 286  
Db 283 GLIAGEQHDLKRMMNGFSKSGKS--YVMAVGVQAVSWQVYVWVGLVGLVSVSVLSNVI 340

QY 287 ISVLLPVTVEFVAVCFR---EKFOAEKGVSLLSLWGFVSYFYGFKSGK 333  
Db 341 SVITWPIVSVLVVIFNFMDDEDFADFGVALVTAVALSAAAYFFRLHKDNR 390  
RESULT 12  
E96516  
Query Match 17.4%; Score 313; DB 2; Length 389;  
Best Local Similarity 28.0%; Pred. No. 1.4e-17;  
Matches 99; Conservative 60; Mismatches 150; Indels 42; Gaps 10;  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: E96516  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,  
Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.  
A;Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; Tallon,  
ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.  
A;Title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.  
A;Reference number: A86141; MUID:21016719; PMID:11130712  
A;Accession: E96516  
A;Status: preliminary  
A;Molecule type: DNA  
A;Residues: 1-389 <STO>  
A;Cross-references: GB:AE005173; NID:g5668802; PIDN:AAD46028.1; GSPDB:GN00141  
C;Genetics:  
A;Gene: Fl6N3.13  
A;Map position: 1

QY 3 NGIIINCILLITGCGPPLRLRYLTNGG-----KRIWFMFSLTAGFP-IILI 51  
Db 52 SGFVTVGRVLST-----LLNNYFIQTGRDADDPKQFGKGTWLSQFLQNAFPSTAF 104  
QY 52 PLLVSLFSLRRGRNPNNAENKRTKLFMETPLFIASIVIGLTLGLDNLVSLYGLAYLP 111  
Db 105 LLLWRSLSFTHGETQSSTSGFK-----LFLYISLGLVSLFSAYSOLYAGRTHCV 153  
QY 112 VSTSSLIIGTQLAFNALPAFLVKKQFTPFPSINAVLLTVIGIGILALHSDGDKPAKESK 171  
Db 154 F--FFWIFTQLIFTIFITAINKHKNRWIILSVLGVATGITS--SDDAYYPCESEG 210  
QY 172 EYVVGFLMTVVAALYAFILPLVELTYKKA--RQETTFPLVLEIQNMVCLAAATFFCVIGM 229  
Db 211 KMSYGAWCSFFGTVAFLSLCINQLGQKQVPIKTESRVSAMVLMOTNASMIATLICLVGL 270  
QY 230 FIVGDPKVIAREAREFKIGSVFPYVALIVITGIIWQGFILGAIGIVFCASSLASGLVLSV 289  
Db 271 FVSGEFDKIDEPETKTGPKLYVLSLIGLS-LAWQVMSLGLVGLVCLASSLSFNVSF 329  
QY 290 LLPVTEFVAVCFREKFOAE-----KGVSLLSLWGFVSYFYGFKSGK 334  
Db 330 ATPVANIFVVLAPR--FMDDIGWFKGALLAGILGFASVYSLYKAIRK 377  
RESULT 13  
D96613  
Query Match 18.2%; Score 327; DB 2; Length 393;  
Best Local Similarity 26.6%; Pred. No. 1.1e-18;  
Matches 93; Conservative 68; Mismatches 159; Indels 30; Gaps 8;  
C;Species: Arabidopsis thaliana (mouse-ear cress)  
C;Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001  
C;Accession: D96613  
R;Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Alonso,  
Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Huizar, L.; Dewar, K.;  
ansen, N.F.; Hughes, B.; Huizar, L.  
Nature 408, 816-820, 2000  
A;Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; Kim, C.;  
C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, Z.A.; Luros, J.S.; Maiti, R.; Marziali,

QY 5 LIINICILITICGCGPLLIRLYFTNGGKRI-----WFMSPSLSTAGPFIILIPLL 54

Db 48 IFVCSCLVA-GRVLSTLLNFYFQTGDEVCDPKQKGTWLOQSVQNAAPFTAFLL 106

QY 55 V---SFLSRRRRNRNPNAENKRTKFLMETPLFIASIVIGLLTGLDNLVSYG---L 107

Db 107 LWRSSF-----STHSTSSSSSF---GKLFLLYISLGLFAAYSQLVAIRTHCV 154

QY 108 AYLPVSTSLIIGTOLAFNALFAFLLVKQKTFPESINAVLLTVGIGILALHSDGDK--P 165

Db 155 FFLWIFTS-----QLIFSIPTTIINKQKFNRIILSWLWSGAATG-LGITSGGAYIP 207

QY 166 AKESKEYVVGFLMTVVAALLVAFILPVELVYKKA--RQETFPVLVLEIQVMVCLAA 223

Db 208 CENEGSKNSGAWCAFFGIVAFSLSLCINQLGFKQKVIPTTQSRVSAVILMQTNASMIATL 267

QY 224 FCVIGMFIVGDPKVIAREAREPKIGGSVPYVALIVITGIWQGFAGIVFCASLAS 283

Db 268 ICVLGLFVSGEFKDIEDFETFKGKPLVLSLIGLS-LAWQVMSLGLVLGLVCLASLFS 326

QY 284 GVLISVLLPVTEVFAVVCPR-----EKFOAEKGVSLLLSLWGFVSVPYGFKSGKK 334

Db 327 NVVSPCSTPLNVILLVAFRTDADVKFKEG--ALVAGILGFASYVVSILYKSTKK 380

RESULT 15

C96613

hypothetical protein F313D3.4 [imported] - Arabidopsis thaliana

C:Species: Arabidopsis thaliana (mouse-ear cress)

C:Date: 02-Mar-2001 #sequence\_revision 02-Mar-2001 #text\_change 31-Mar-2001

C:Accession: C96613

R:Theologis, A.; Ecker, J.R.; Palm, C.J.; Federspiel, N.A.; Kaul, S.; White, O.; Chin, C.W.; Chung, M.K.; Conn, L.; Conway, A.B.; Conway, A.R.; Creasy, T.H.; D.ansen, N.F.; Hughes, B.; Huizlar, L.

Nature 408, 816-820, 2000

A:Authors: Hunter, J.L.; Jenkins, J.; Johnson-Hopson, C.; Khan, S.; Khaykin, E.; C.A.; Li, J.H.; Li, Y.; Lin, X.; Liu, S.X.; Liu, Z.A.; Lueros, J.S.; Maiti, R.; Rizzo, M.; Rooney, T.; Rowley, D.; Sakano, H.

A:Authors: Salzberg, S.L.; Schwartz, J.R.; Shinn, P.; Southwick, A.M.; Sun, H.; ker, M.; Wu, D.; Yu, G.; Fraser, C.M.; Venter, J.C.; Davis, R.W.

A:title: Sequence and analysis of chromosome 1 of the plant Arabidopsis.

A:Reference number: A86141; MUID:21016719; PMID:11130712

A:Accession: C96613

A>Status: preliminary

A:Molecule type: DNA

A:Residues: 1-398 <STO>

A:Cross-references: GB:AE005173; NID:gl1128382; PIDN:AAG31188.1; GSPDB:GN00141

C:Genetics:

A:Gene: F313D3.4

A:Map position: 1

Query Match 15.1%; Score 271; DB 2; Length 398;

Best Local Similarity 25.2%; Pred. No. 3.3e-14;

Matches 82; Conservative 65; Mismatches 157; Indels 22; Gaps 6

QY 33 KRIWMSFLSTAGPPIILIPLLVSFLSRRNRNPNAENKRTKFLMETPLFIASIVI 92

Db 85 KGTWTQALIQNAAPP-ILILPFFIF-----PKPKHLETNNTSLSRLRFLPLYLSL 134

QY 93 GLTLGLDNLVSYGLAYLPVSTSLIIGTOLAFNALFAFLLVKQKTFPESINAVLLTVG 152

Db 135 GVLVAHSLKFLPALKLVSNVYGIPLSIITQLIFAVLTAINRFKTRWIIIS-ILLTIV 193

QY 153 IGILALHSDGDKPAKESKEYVVGFLMTVVAALLYAFILPVELVYKKA-----QEI 205

Db 194 IYVLGTDPFGQPHDGEFGFYNIQAWLAFSATIAFSLSLCIIQLGFEKLQVTKRYGNEK 253

QY 206 TFFLVLEIQVMVCLAAATFFCVIGMFIVGDPKVIAREAREPKIGGSVPYVALIVITGI 265

Db 254 VFRVWLEMQICVAFVSVVCLVGLFASDEYKELGDSKRFK-KGTYYVLSVLGLALSQ 312

QY 266 GFFLGAIGIVFCASSLASGVLISVLLPVTEVFAVVCPR---REKFOAEKGVSLLLSLWGFV 322



Db 313 VWAUGHIGLVHYVSLFGDVVHMCASPFVALFVVLAFDFMDDVFSWFRIGALIGTVLALG 372  
Qy 323 SYFYGEFKSGKVVDRKPPETELPI 348  
Db 373 SYFYTLHKRNKKQWAEINQSENNVEV 398

Search completed: September 23, 2004, 23:51:16  
Job time : 27.1412 secs

**THIS PAGE BLANK (US-12)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:50:27 ; Search time 85.017 Seconds  
(without alignments)  
1331.366 Million cell updates/sec

Title: US-09-913-767-9  
Perfect score: 1760  
Sequence: 1 MKVALVINCIILAIACGG.....QQEESQETQSLSLRPISEC 352

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1349238 seqs, 321558718 residues

Total number of hits satisfying chosen parameters: 1349238

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : Published Applications\_AA:  
1: /cgn2\_6/ptodata/2/pubpaa/US07\_PUBCOMB.pep.\*  
2: /cgn2\_6/ptodata/2/pubpaa/PCT\_NEW\_PUB.pep.\*  
3: /cgn2\_6/ptodata/2/pubpaa/US06\_NEW\_PUB.pep.\*  
4: /cgn2\_6/ptodata/2/pubpaa/US06\_PUBCOMB.pep.\*  
5: /cgn2\_6/ptodata/2/pubpaa/US07\_NEW\_PUB.pep.\*  
6: /cgn2\_6/ptodata/2/pubpaa/PCTUS\_PUBCOMB.pep.\*  
7: /cgn2\_6/ptodata/2/pubpaa/US08\_NEW\_PUB.pep.\*  
8: /cgn2\_6/ptodata/2/pubpaa/US08\_PUBCOMB.pep.\*  
9: /cgn2\_6/ptodata/2/pubpaa/US09A\_PUBCOMB.pep.\*  
10: /cgn2\_6/ptodata/2/pubpaa/US09B\_PUBCOMB.pep.\*  
11: /cgn2\_6/ptodata/2/pubpaa/US09C\_PUBCOMB.pep.\*  
12: /cgn2\_6/ptodata/2/pubpaa/US09\_NEW\_PUB.pep.\*  
13: /cgn2\_6/ptodata/2/pubpaa/US10A\_PUBCOMB.pep.\*  
14: /cgn2\_6/ptodata/2/pubpaa/US10B\_PUBCOMB.pep.\*  
15: /cgn2\_6/ptodata/2/pubpaa/US10C\_PUBCOMB.pep.\*  
16: /cgn2\_6/ptodata/2/pubpaa/US10\_NEW\_PUB.pep.\*  
17: /cgn2\_6/ptodata/2/pubpaa/US60\_NEW\_PUB.pep.\*  
18: /cgn2\_6/ptodata/2/pubpaa/US60\_PUBCOMB.pep.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	Query Match	Length	ID	Description
1	909	51.6	320	12	US-10-424-599-173116 Sequence 173116,
2	866	49.2	390	16	US-10-437-963-160022 Sequence 160022,
3	841.5	47.8	306	12	US-10-424-599-216343 Sequence 216343,
4	830	47.2	293	12	US-10-425-114-44529 Sequence 44529, A
5	619.5	35.2	423	16	US-10-437-963-184544 Sequence 184544,
6	590.5	33.6	391	16	US-10-437-963-193908 Sequence 193908,
7	572.5	32.4	362	12	US-10-424-599-245361 Sequence 245361,
8	552.5	31.5	409	12	US-10-424-599-239885 Sequence 239885,
9	540	30.7	366	12	US-10-424-599-145366 Sequence 145366,
10	534	30.3	751	16	US-10-437-963-146343 Sequence 146343,
11	528	30.0	267	12	US-10-425-114-61519 Sequence 61519, A
12	508.5	28.9	355	12	US-10-424-599-174206 Sequence 174206,
13	500.5	28.4	246	12	US-10-425-114-70607 Sequence 70607, A
14	494.5	28.1	363	12	US-10-424-599-174205 Sequence 174205,
15	492	28.0	372	16	US-10-437-963-136993 Sequence 136993,

16	489.5	27.8	216	16	US-10-437-963-174935 Sequence 174935,
17	489.5	27.8	379	12	US-10-425-114-45872 Sequence 45872, A
18	476.5	27.1	400	16	US-10-437-963-104380 Sequence 104380,
19	476	27.0	373	16	US-10-437-963-144907 Sequence 144907,
20	472.5	26.8	610	16	US-10-437-963-186383 Sequence 186383,
21	471	26.8	379	16	US-10-437-963-121341 Sequence 121341,
22	441.5	25.1	327	12	US-10-424-599-218003 Sequence 218003,
23	407.5	23.2	238	12	US-10-424-599-204152 Sequence 204152,
24	343	19.5	241	12	US-10-424-599-152055 Sequence 152055,
25	294	16.7	202	16	US-10-767-701-38370 Sequence 38370, A
26	270.5	15.4	164	16	US-10-767-701-40564 Sequence 40564, A
27	235.5	13.4	166	16	US-10-767-701-33609 Sequence 33609, A
28	228.5	13.0	173	12	US-10-424-599-257773 Sequence 257773,
29	216.5	12.3	70	12	US-10-424-599-261293 Sequence 261293,
30	202	11.5	167	16	US-10-767-701-35590 Sequence 35590, A
31	194.5	11.1	150	16	US-10-767-701-37550 Sequence 37550, A
32	167.5	9.5	94	12	US-10-424-599-263994 Sequence 263994,
33	156.5	8.9	357	16	US-10-437-963-151972 Sequence 151972,
34	138.5	7.9	358	16	US-10-437-963-186107 Sequence 186107,
35	136.5	7.8	388	12	US-10-425-114-44759 Sequence 44759, A
36	134.5	7.6	296	16	US-10-767-701-35041 Sequence 35041, A
37	133.5	7.6	314	12	US-10-282-122A-53153 Sequence 53153, A
38	132	7.5	102	12	US-10-424-599-217019 Sequence 217019,
39	131	7.4	327	12	US-10-424-599-271242 Sequence 271242,
40	130.5	7.4	380	16	US-10-437-963-106742 Sequence 106742,
41	129.5	7.4	369	12	US-10-425-114-64886 Sequence 64886, A
42	129.5	7.4	592	16	US-10-437-963-147143 Sequence 147143,
43	129	7.3	296	9	US-09-815-242-5098 Sequence 5098, Ap
44	129	7.3	296	12	US-10-282-122A-43609 Sequence 43609, A
45	128.5	7.3	313	12	US-10-425-114-72516 Sequence 72516, A

ALIGNMENTS

RESULT 1

US-10-424-599-173116  
; Sequence 173116, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 173116  
; LENGTH: 320  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_12733C.1.pap  
US-10-424-599-173116

Query Match 51.6%; Score 909; DB 12; Length 320;  
Best Local Similarity 55.2%; Pred. No. 3,7e-81;  
Matches 170; Conservative 66; Mismatches 68; Indels 4; Gaps 2;

QY	24	MRLYFNNGGKRIWFSTFLETAGPPVIFIPLLPSYITRRSNVVDSTSPFLIPRLIIA 83
DB	1	VRDYFHGGKRIWLSFLETAANSARVLPIIISYIKRRHRH---SSALISIKPPLFLAS 57
QY	84	VIVGILSGFDNLYAYGIAYLPVSTAALIIASQALIAIFSPFMVXKHTFPTINAVLL 143
DB	58	ALIGLTLGDDLYIACGVARLPVSTFSLQASHLATVAFLLVHRHRTFPYVSNVLL 117
QY	144	TGGAAYLGMHTETDKPVHETHKQYITGFLITVAAAVMYAFILPLVLAYOKAKQMTSYTL 203
DB	118	TVAAVVLAALRSSGDRPAGESRQYVIGFVMIILAAAALYGVFLPLMELVYKKSQRITSL 177

```
QY 204 VLEFOLILCLLASIVSVIGMFIAGDFKQALPKAREFKLGEALFYVWVAFSAIIWQGFLL 263
Db 178 VNEIQVLVLCFATLPCITVGMIIINDDFK-VIPREARDFKLGETKYVVLVWSAIIWQGFLL 236
QY 264 GAIGLIFSTSSIVSGIMISVLLPTEVLAVFYHKKFOABKGLSLALSLWGFVSFYFGEI 323
Db 237 GAIGVIFCASSLLSIIIAAFLPVTVEVLAVIYKESFPAEKGVALVLSLWGFVSFYFGEI 296
QY 324 KSGEDKRR 331
Db 297 KODREKNK 304
```

```
RESULT 2
US-10-437-963-160022
; Sequence 160022, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 160022
; LENGTH: 390
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_59340C.1.pap
US-10-437-963-160022
```

```
Query Match 49.2%; Score 866; DB 16; Length 390;
Best Local Similarity 47.6%; Pred. No. 8.4e-77;
Matches 161; Conservative 85; Mismatches 80; Indels 12; Gaps 3;

QY 5 LVINICIIAIGNCGGFLIMRLYFNNGGKRIFWFTFLETAGFPVIFIPLLFSYITRRSN 64
Db 39 LNVNVLVWVGSACGFLLLRAYFLRGNGKWLSSLLQTAGWPLLALPCFSYSSRRRR 98
QY 65 NVGD-----STSFLLKPRLLIAAVIVGILSGFDNLYAYGIAYLPVSTAALIISQ 116
Db 99 EVEDDGAGAGAAATPLFMTPELLVASAVVGLMTGVDLLIAYGLAYLPVSTSSILISIQ 158
QY 117 LAFIAIFPFMVKHKTPTTINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVA 176
Db 159 LATAPALLVQRFTASVNAVLLSVGAAMLGNAGGDRPAGVSRQAQYAGFAMTLA 218
QY 177 AAVMYAFILPLVELA---YOKAKQTSYTLVLEFQILCLLASIVSVIGMFIAGDFKQAL 233
Db 219 AAALYGLVLPVWELSQAHAAARGAVTYTLVMEMQLVIGFVATFSAVGLVNNDF-HAI 277
QY 234 PKAREPKLGEALFYVWVAFSAIIWQGFGLGIFSTSSIVSGIMISVLLPTEVLAV 293
Db 278 PGSAHEFGLQAGYLLLAGSAAAMYOCFFLGTIGALFYGSALLAGVIMTILPVTVELAV 337
QY 294 IFYHKKFOABKGLSLALSLWGFVSFYFGEIKSGEDKRR 331
Db 338 MFPEHPNGTKGVALALSLWGFVSFYFGEVRAAKAAHR 375
```

```
RESULT 3
US-10-424-599-216343
; Sequence 216343, Application US/10424599
```

```
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 216343
; LENGTH: 306
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(306)
; OTHER INFORMATION: unsure at all xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_37387C.1.pap
US-10-424-599-216343
```

```
Query Match 47.8%; Score 841.5; DB 12; Length 306;
Best Local Similarity 54.6%; Pred. No. 1.6e-74;
Matches 160; Conservative 59; Mismatches 71; Indels 3; Gaps 2;

QY 1 MKVALVIINCIIAIGNCGGFLIMRLYFNNGGKRIFWFTFLETAGFPVIFIPLLFSYITR 60
Db 15 MKRLLLTNCLLTIGTSGGFLVMRLYFLRGHVRVLSSELETAGFPMLPLPLAVSYLR 74
QY 61 RRSNNVGDSTSFELJ---KPRLLIAAVIVGILSGFDNLYAYGIAYLPVSTAALIISQLA 118
Db 75 RRTASAGATAPKLUISMKPPLLAFTFILTGLDDVLYAYGVARLPVTSALIIATQLG 134
QY 119 FIAIFSPFMVKHKTPTTINAVLLTVGAALVGMHTETDKPVHETHKQYITGFLITVAAA 178
Db 135 FTAFFAFLVQRKFTAYSVNAVLLTVGAGVLAALHTSGDRPPGESVKEYVMGFVMTVIAA 194
QY 179 VMYAFILPLVELAYOKAKQTSYTLVLEFQILCLLASIVSVIGMFIAGDFKQALPKEAR 238
Db 195 ALYGFILPLVELVYKIKQPLTVSLVMEIOFVMCFSATLFLGLGMIINDDFK-VIPREAK 253
QY 239 EFKLGEALFYVWVAFSAIIWQGFGLGIFSTSSIVSGIMISVLLPTEVL 291
Db 254 KFEHGESYAYLVGSAILWQAFLGAIGVIFWAWLLIXGFLIAVLLPVTVEL 306
```

```
RESULT 4
US-10-425-114-44529
; Sequence 44529, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 44529
; LENGTH: 293
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: 700941791_FLI.pap
US-10-425-114-44529
```

Query Match 47.2%; Score 830; DB 12; Length 293;  
Best Local Similarity 54.8%; Pred. No. 2.1e-73;  
Matches 154; Conservative 63; Mismatches 60; Indels 4; Gaps 2;  
QY 51 IPLLFSYITRRSNVYVGGDSTFPLIKPRLLIAAVIVGILSGFNDYLYAYGIAYLPVSTAA 110  
DB 1 LPIIISYIHKRRHR---SALISIKPPLFLASALIGLTLGLDYLACVAVLPVSTFS 57  
QY 111 LIASOLAFIAIPFPMVKHKPTFTINAVLLTVGAALVGMHETDKPVHETHKQVITG 170  
DB 58 LIQASHLATVAFALVLRHFTFYSVNSVLLTVAAVLAALRSSGDRPAGESRQYVIG 117  
QY 171 FLITVAAAVYAFILPLVELAYOKAKTMSYTYLVLEFQLILCLLASIVSVIGMFIAGDFK 230  
DB 118 FVMILAAAALYGFVPLMELVYKSRQITYSLVMEITQLVLCFFATLCTVGMINDDFK 177  
QY 231 QALPKAREFKLGEALFYVAVFSAIIWQGFGLGALIFSTSLVSGIMISVLLPITEV 290  
DB 178 -VIPREARDFKGETKYIVVVLVWSAIMWQFFFGAIGVIFCASLLSGIIIAAFLPVTEV 236  
QY 291 LAVIFYHEKFOAKGSLALSLSLWGFVSFYFYGEIKSGEDKRR 331  
DB 237 LAVIVYKESFHAKEGVALVLSLWGFVSFYFYGEIKQDRKNK 277

RESULT 5  
US-10-437-963-184544  
; Sequence 184544, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT FILING DATE: 2003-05-14  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 184544  
; LENGTH: 423  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; LOCATION: (1)..(391)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_81525C.1.pep  
US-10-437-963-184544

Query Match 35.2%; Score 619.5; DB 16; Length 423;  
Best Local Similarity 36.8%; Pred. No. 2.1e-52;  
Matches 123; Conservative 84; Mismatches 124; Indels 5; Gaps 3;  
QY 5 LVINCIIILAIAGNCGGLIMRLVFNNGKRIWFTSTLETAGRPVIFIPILLFSYITRRSN 64  
DB 89 MGVNMFLLIAGQTASTLGRFYINGGNSKMWSTFVQTAGFPILLFALFLPHKTSSTQ 148  
QY 65 NVGDSTSFLLIKPRLLIAAVIVGILSGFNDYLYAYGIAYLPVSTAAIIASQLAFIAIF 124  
DB 149 TVTSSPAPITISPKITLIYVVLGLIATAADLMYSYGLLYLPVSTYSLICASQLAFNAVS 208  
QY 125 FPMVKHKPTFTINAVLLTVGAALVGMHETDKPVHETHKQVITGFLITVAAAVYAFI 184  
DB 209 YFLNAQKQFLIFNSVLLTFSSALILGVDEDSQGTTSISHGKYLGLLTLGASATYSLI 268  
QY 185 LPILVELAYOKAKTMSYTYLVLEFQLILCLLASIVSVIGMFIAGDFKQALPKAREFKLGE 244  
DB 269 LSLMQVTFEKKVIRETFSSVVLNMQIYALVATLASLVGLFASGEW-MTLQGMHFQSGK 327

QY 245 ALFYVAVFSAIIWQGFGLGALIFSTSLVSGIMISVLLPITEVLAVIFYHEKFOAEK 304  
DB 328 LSVVMTLLTAISQVAVSGVGLIFWSSLSFNSVISTLALPIPVFAVIFHDKMDGVK 387  
QY 305 GLSLALSLSLWGFVSFYFYGEIKSGEDKRIQOERS 337  
DB 388 ITAMLMAIWGFMSYGHQLYVDGKKGR-KTTVSVEET 422

RESULT 6  
US-10-437-963-193908  
; Sequence 193908, Application US/10437963  
; Publication No. US20040123343A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa, Thomas J.  
; APPLICANT: Kovalic, David K.  
; APPLICANT: Zhou, Yihua  
; APPLICANT: Cao, Yongwei  
; APPLICANT: Wu, Wei  
; APPLICANT: Boukharov, Andrey A.  
; APPLICANT: Barbazuk, Brad  
; APPLICANT: Li, Ping  
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With  
; FILE REFERENCE: 38-21(53221)B  
; CURRENT FILING DATE: 2003-05-14  
; CURRENT APPLICATION NUMBER: US/10/437,963  
; NUMBER OF SEQ ID NOS: 204966  
; SEQ ID NO 193908  
; LENGTH: 391  
; TYPE: PRT  
; ORGANISM: Oryza sativa  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(391)  
; OTHER INFORMATION: unsure at all Xaa locations  
; OTHER INFORMATION: Clone ID: PAT\_MRT4530\_90000C.1.pep  
US-10-437-963-193908

Query Match 33.6%; Score 590.5; DB 16; Length 391;  
Best Local Similarity 39.0%; Pred. No. 1.4e-49;  
Matches 137; Conservative 67; Mismatches 110; Indels 37; Gaps 8;  
QY 5 LVINCIIILAIAGNCGGLIMRLVFNNGKRIWFTSTLETAGRPVIFIPILLFSYITRRSN 64  
DB 49 LVIFSACLVLG-AGGPLLLRVFVHGGRLLWSATLQISGWPELLLPCLCVS-LYGRRH 106  
QY 65 NVGDSTSFLLIKPRLLI-AAVIVGILSGFNDYLYAYGIAYLPVSTAAIIASQLAFIAIF 123  
DB 107 GIGN-----LLLPRLVGAALVGLYAVSCFVVALGSQALPLSTSSLLATQATPAVF 161  
QY 124 SPMVKHKPTFTINAVLLTVGAALVGMHETDKPVHETHKQVITGFLITVAAAVYAF 183  
DB 162 AFLFLGRFTFSSANAVLLTVRVRGRCKTVDRPDHGGWHVD----- 208  
QY 184 ILPLVELAYQ-----KAKQTSYTYLVLEFQLILCLLASIVSVIGMFIAGDFK 230  
DB 209 VQOVTVLQSDWDMCEGVLGGELPPTSPYATVMQAVMGADTXCVGLGNAIGAF- 267  
QY 231 QALPKAREFKLGEALFYVAVFSAIIWQGFGLGALIFSTSLVSGIMISVLLPITEV 290  
DB 268 QAVAREAAFGLAANYLVLAWDVSMQLNLINGLITCASLLAGIMIAVLLPLSQV 327  
QY 291 LAVIFYHEKFOAKGSLALSLSLWGFVSFYFYGEIKSGEDKRIQOERSQETE 341  
DB 328 LAVIFLHEKFDGTGKIALVLSLWGFASLYGE--KAKKKEAKQMREREQE 376

RESULT 7  
US-10-424-599-245361  
; Sequence 245361, Application US/10424599  
; Publication No. US20040031072A1

GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 245361  
; LENGTH: 362  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure  
; LOCATION: (1)..(362)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_63592C.1.pap  
US-10-424-599-245361

Query Match 32.5%; Score 572.5; DB 12; Length 362;  
Best Local Similarity 36.3%; Pred. No. 7.4e-48;  
Matches 123; Conservative 72; Mismatches 137; Indels 7; Gaps 3;

QY 5 LVIINCIIAIGNCGPLIMRLYFNNGKRIWFSTFLETAGPPVIFIPILLFSYTRRRSN 64  
Db 30 LVTLSTAFIVGCSAAVLGRFYDGGNSKMWATLVQTAAPFIFLFTIPSPPEAST 89  
QY 65 NVGDSSTFELIKPRLLIAAVIVGILSGFDNYLYAGIAYLPVSTAAALIASOLAFIIFS 124  
Db 90 SASPP-----IKILLIYXVF-GVLIADNMYSGLLYLSASTYSLICASOLAFNAFVS 143  
QY 125 FPMVKHFTPTTINAVLLTVGAALVGMHTETDKPVHETHKQYITGLITVAAAVMYAFI 184  
Db 144 YFINSQKFTALINSTVLTSAALLAVNEDTDESGFGSKYIIIGFLCTIGASAVYSL 203  
QY 185 LPVLAELAYOKAKOTMSYTLVLEFQLICLLASIVGIMFAGDPKQALPKAREEFKGE 244  
Db 204 LSLMQJTFEKLKETFVSIVLEQIYTSFVASCASVIGLFASGEWR-TLHGEMEGFQKGY 262  
QY 245 ALFYVAVFSAITWQGFFLGAIGLIPSTSSVSGIMISVLLPITEVLAVIFHEKFAEK 304  
Db 263 VAVVMTLWTSIAWQVCSGVGGLIFLVSLSYNSVISTVSLAVTPTAAVIVFHDKMGVYK 322  
QY 305 GLSLALSLGCFVSFYFGEIKSGEDKRIQOEBESQETEQS 343  
Db 323 IISMLLALMGFASYIYQNYLDLDSKTRHAQAATKSQNDSS 361

RESULT 8  
US-10-424-599-239885  
; Sequence 239885, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 239885  
; LENGTH: 409  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; NAME/KEY: unsure

Query Match 30.7%; Score 540; DB 12; Length 366;  
Best Local Similarity 34.8%; Pred. No. 1.2e-44;  
Matches 123; Conservative 80; Mismatches 114; Indels 36; Gaps 9;

QY 18 CGGPLIM--RLYFNNGKRIWFSTFLETAGPPVIFIPILLFSYTRRRSNVGDSTSFLLI 75  
Db 28 CGSAAILLGRLYEYKGGKSMGTVLQVAGFP--IQPFHFIS--ASKNUTTSS---I 79  
QY 76 KPRLLIAAVI-----VGLSGFDNYLYAGIAYLPVSTAAALIASOLAFIIFSFFVMK 129  
Db 80 HPKQSASALAFIYVSGILLLALDCLYISVGLWDLPVSTYSLICSSQLAFNAFFSYFLNS 139  
QY 130 HKFTPTTINAVLLTVGAALVGMHTETDKPVHETHKQYITGLITVAAAVMYA 182  
Db 140 LKFTPTIINSLVLTISSTLLVFNQNESSDDDDSDSTKISKKKYVIGICTVGASAGY 199

LOCATION: (1)..(409)  
; OTHER INFORMATION: unsure at all Xaa locations  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_58641C.1.pap  
US-10-424-599-239885

Query Match 31.4%; Score 552.5; DB 12; Length 409;  
Best Local Similarity 34.1%; Pred. No. 8.2e-46;  
Matches 118; Conservative 79; Mismatches 132; Indels 17; Gaps 5;

QY 5 LVIINCIIAIGNCGPLIMRLYFNNGKRIWFSTFLETAGPPVIFIPILLFSYI---TRR 61  
Db 66 LLVINYACLFGVSLSSLLSKYVFTHKGSSRWSTWVQTAGPFLIPICVPLFKFTKR 125  
QY 62 RSNVGDSTSFELIKPRLLIAAVIVGILSGFDNYLYAGIAYLPVSTAAALIASOLAFIA 121  
Db 126 VPFN--DFT-----PRMLIISIGVMGLFNLLPFSWGNISYLPVSTSAALLSSQLLFWL 177  
QY 122 IFSFPMVKHFTPTTINAVLLTVGAALVGMHTETDKPVHETHKQYITGLITVAAAVMY 181  
Db 178 LFSVIVKQKIIFSNVNCVILLTSSILIALDSSHERPKGLTKNYFIFGFFCTIGAGLMF 237  
QY 182 AFILPLVELAYOKAKOTMSYTLVLEFQLICLLASIVGIMFAGDPKQALPKAREEFK 241  
Db 238 ALVPLMEKIYKVN--CYQVMEMQXIMEGAATAIAIVGMTWGGFSEMKVESQMVFD 294  
QY 242 LGHALFYVAVFSAITWQGFFLGAIGLIPSTSSVSGIMISVLLPITEVLAVIFHEKFO 301  
Db 295 KGSRVVYVMTVNGVNVWQLCFMGTAGWVFLTSSLTGICMTWFLSNVNLGGVVFFRDAFG 354  
QY 302 AEGLSLALSLGCFVSFYFGEIKS---GEDKRIQOEBESQETEQS 344  
Db 355 GVKAVSTFCLGFCSVYGIYKQNGQNGEHKLASTRNTKTISSNASS 400

RESULT 9  
US-10-424-599-145366  
; Sequence 145366, Application US/10424599  
; Publication No. US20040031072A1  
; GENERAL INFORMATION:  
; APPLICANT: La Rosa Thomas J  
; APPLICANT: Kovalic David K  
; APPLICANT: Zhou Yihua  
; APPLICANT: Cao Yongwei  
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With  
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement  
; FILE REFERENCE: 38-21(53223)B  
; CURRENT APPLICATION NUMBER: US/10/424,599  
; CURRENT FILING DATE: 2003-04-28  
; NUMBER OF SEQ ID NOS: 285684  
; SEQ ID NO 145366  
; LENGTH: 366  
; TYPE: PRT  
; ORGANISM: Glycine max  
; FEATURE:  
; OTHER INFORMATION: Clone ID: PAT\_MRT3847\_102285C.1.pap  
US-10-424-599-145366

Query Match 30.7%; Score 540; DB 12; Length 366;  
Best Local Similarity 34.8%; Pred. No. 1.2e-44;  
Matches 123; Conservative 80; Mismatches 114; Indels 36; Gaps 9;

QY 18 CGGPLIM--RLYFNNGKRIWFSTFLETAGPPVIFIPILLFSYTRRRSNVGDSTSFLLI 75  
Db 28 CGSAAILLGRLYEYKGGKSMGTVLQVAGFP--IQPFHFIS--ASKNUTTSS---I 79  
QY 76 KPRLLIAAVI-----VGLSGFDNYLYAGIAYLPVSTAAALIASOLAFIIFSFFVMK 129  
Db 80 HPKQSASALAFIYVSGILLLALDCLYISVGLWDLPVSTYSLICSSQLAFNAFFSYFLNS 139  
QY 130 HKFTPTTINAVLLTVGAALVGMHTETDKPVHETHKQYITGLITVAAAVMYA 182  
Db 140 LKFTPTIINSLVLTISSTLLVFNQNESSDDDDSDSTKISKKKYVIGICTVGASAGY 199

```

183 QY 183 FILPIVELAYOKAKQMTGMYTIVLEFQILCLLLASIVSVGMFTAGDFKQALPKEARBFKL 242
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
200 Db 200 LMLSITQLVFKKIKRETFFKVQDLMDILYTSLVATLATLVGLFASGEW-SGLKNEMKEYEL 258
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 243 GEALFYVVAVFSALIIWQGFGLGALIFSTSSIVSGIMISVLLPITEVLAVIFYHEKFOA 302
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 259 GKASYLLMLTTFALWQVFTTICGLGDISEVSSLFSNAISALGVPIVPMPLAVLFFHDKMDG 318
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 303 EKGSLALSMLGWFVSFYFYGEIKSGEDKRRIQEESQETEQSS-----LSRPISE 351
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db 319 IKGISMVLAINGIVSYVY-----QQVLDDTKSENRRNTTSHVPKASSPIEE 363
    : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :

RESULT 10
US-10-437-963-146343
; Sequence 146343, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 146343
; LENGTH: 751
; TYPE: PRT
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_46977C.1.1.pep
US-10-437-963-146343

```

```

; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaeka, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 61519
; LENGTH: 267
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: LIB143-024-D8_FLI.pep
; US-10-425-114-61519

Query Match          30.0%; Score 528; DB 12; Length 267;
Best Local Similarity 43.8%; Pred. No. 1.2e-43;
Matches 102; Conservative 55; Mismatches 72; Indels 4; Gaps 2;

Qy 1 MVKALVIINCIIAIGNCGPLIMRIYFNNGKRIWPFSTLETAGPVPVIFIPLLFSYITR 60
Db 26 MRRLVATNCVMLALGVTCQGLLSRLYFSKGGHROWLSGWLQTCGWPLLLPPVAASYV-R 84

Qy 61 RRSNNVGDSTSFPLIKPRLLIAAVIVGILSGEDNYLYAYGIALVPVSTRAALIISOLAFI 120
Db 85 RRARDSAPVLLTQTQPRILLAAAGLIGAGVDNLLYAMGLEFLFPVSTSAIISTQLAFT 144

Qy 121 AIFSFFWVKHKPTPTTINAVLLTVGAAVLGMHTETDKPVHETHKQVITGFLITVAAAVM 180
Db 145 VLFAYLIVRQLRTMATVNAVLLTVGANVGLHVSSDRPAGVTRSQYWLGFILTLGAAVL 204

Qy 181 YAFILPVELAYQKAK---QTMSTYIVLVBQLILCLLASIVSVIGMFIAGDFK 230
Db 205 YGLFLPLVELTYKCAAGGGRPVTYALVWELQLVMGFVATACFTVGMIVNKDFQ 257

RESULT 12
US-10-424-599-174206
; Sequence 174206, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa Thomas J
; APPLICANT: Kovalic David K
; APPLICANT: Zhou Yihua
; APPLICANT: Cao Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 174206
; LENGTH: 355
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_128327C.1.pep
; US-10-424-599-174206

```

RESULT 11  
US-10-425-114-61519  
; Sequence 61519, Application US/10425114

```
QY 62 RSNVGDSTFLLKPRLLIAAVIVGILSGFDNYLXAYGIAYLPVSTAALIASQLAFIA 121
Db 102 TSLNL-----KLSUYIVIGFSLAADNLMYAYAYLPASTASLVASSLVFSA 150
QY 122 IFSFPMVKHFTPTINAVLLTVGAAGVLMGHTETDKPVHETHKQYITGFLITVAAMVY 181
Db 151 LFGYFLVKNKNVSNVSVFIITIALTDSSDRYANISDSEYIMGFVMDVLSALH 210
QY 182 AFILPLVELAYOKAQMTSYTLVLEFQLILCLIASIVSGMISVLLPITEVLAVFYHEKFK 241
Db 211 GLIFALSELVFKLLERSFVLEQQVMVSLFAFLFTTGVGIMSGDF-QQMAHEATTF 269
QY 242 LGEALFYVAVFSAIIMQGFGLGALIFSTSSLVSGMISVLLPITEVLAVFYHEKFK 301
Db 270 GRSAYLVIWGAITFQGLVGGTAVIFLGSVTLVAGVNAVRTPTITSIAAVILLKDPMS 329
QY 302 AEKGLSLALSNGVSYFYGEIK 324
Db 330 GFKLSLVITFWGFGSYTGSSK 352

RESULT 13
US-10-425-114-70607
; Sequence 70607, Application US/10425114
; Publication No. US20040034888A1
; GENERAL INFORMATION:
; APPLICANT: Liu, Jingdong
; APPLICANT: Zhou, Yihua
; APPLICANT: Kovalic, David K.
; APPLICANT: Screen, Steven E
; APPLICANT: Tabaska, Jack E
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53313)B
; CURRENT APPLICATION NUMBER: US/10/425,114
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 73128
; SEQ ID NO 70607
; LENGTH: 246
; TYPE: PRT
; ORGANISM: Zea mays
; FEATURE:
; OTHER INFORMATION: Clone ID: L1B3608-022-E3_FLI.pep
US-10-425-114-70607

Query Match 28.4%; Score 500.5; DB 12; Length 246;
Best Local Similarity 47.7%; Pred. No. 5.7e-41;
Matches 93; Conservative 45; Mismatches 54; Indels 3; Gaps 1;

QY 1 MVKALVIINCIIAIGNCGPLIMRLYFNNGKRIWFSTFLETAGPPVIFILLFSYTR 60
Db 52 MHRFLVALNCGMLAVGAIGPLLSRLYFSGKHROWLSAWLETGGWPLLVIPVAAAFVAR 111
QY 61 RSNVGDSTFLLKPRLLIAAVIVGILSGFDNYLXAYGIAYLPVSTAALIASQLAFI 120
Db 112 RARDR---GAPVLLAPRLLIAAGLVATGADDFIYAGLALPVTSAIISTQLAFT 168
QY 121 AIFSFFMVKHFTPTINAVLLTVGAAGVLMGHTETDKPVHETHKQYITGFLITVAAMV 180
Db 169 VFAFLVVRQLTAASVNAVALLTGVAGVVLGLHVSDDRPVTRGVWLGFLVTLGAAL 228
QY 181 YAFILPLVELAYOKA 195
Db 229 YGLVPLVELAYRRA 243

RESULT 14
US-10-424-599-174205
; Sequence 174205, Application US/10424599
; Publication No. US20040031072A1
; GENERAL INFORMATION:
```

```
; APPLICANT: La Rosa, Thomas J
; APPLICANT: Kovalic, David K
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; TITLE OF INVENTION: Soy Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53223)B
; CURRENT APPLICATION NUMBER: US/10/424,599
; CURRENT FILING DATE: 2003-04-28
; NUMBER OF SEQ ID NOS: 285684
; SEQ ID NO 174205
; LENGTH: 363
; TYPE: PRT
; ORGANISM: Glycine max
; FEATURE:
; NAME/KEY: unsure
; LOCATION: (1)..(363)
; OTHER INFORMATION: unsure at all Xaa locations
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT3847_128326C.1.pep
US-10-424-599-174205

Query Match 28.1%; Score 494.5; DB 12; Length 363;
Best Local Similarity 31.6%; Pred. No. 3.7e-40;
Matches 101; Conservative 81; Mismatches 123; Indels 15; Gaps 3;

QY 5 LVIINCIIAIGNCGPLIMRLYFNNGKRIWFSTFLETAGPP---VIFIPLLFSYTRR 61
Db 48 LVLISILAMLVAFPASSILSRVYDNGGQSKMISWVAVAGWPLTALILFPVVFISKTTP 107
QY 62 RSNVGDSTFLLKPRLLIAAVIVGILSGFDNYLXAYGIAYLPVSTAALIASQLAFIA 121
Db 108 TPLNL-----KLSLSYIVLGLSAADNLMYAYAYLPASTASLVASSLVFSA 156
QY 122 IFSFPMVKHFTPTINAVLLTVGAAGVLMGHTETDKPVHETHKQYITGFLITVAAMVY 181
Db 157 LFGYFLVKNKNVSNVSVFIITIALTDSSDRYANISDSEYIMGFVMDVXGSAPH 216
QY 182 AFILPLVELAYOKAQMTSYTLVLEFQLILCLIASIVSGMISVLLPITEVLAVFYHEKFK 241
Db 217 GLIFALSELVFKLLGRSFIVVLEQQVMVSLFAFLFTTGVGIMVSGDF-QQMAHEATTF 275
QY 242 LGEALFYVAVFSAIIMQGFGLGALIFSTSSLVSGMISVLLPITEVLAVFYHEKFK 301
Db 276 GRSAYLVIWGAITFQGLVGGTAVIFLGSVTLVAGVNAVRTPTITSIAAVILLKDPMS 335
QY 302 AEKGLSLALSNGVSYFYG 321
Db 336 GFKLSLVITFWGFGSYIYG 355

RESULT 15
US-10-437-963-136993
; Sequence 136993, Application US/10437963
; Publication No. US20040123343A1
; GENERAL INFORMATION:
; APPLICANT: La Rosa, Thomas J.
; APPLICANT: Kovalic, David K.
; APPLICANT: Zhou, Yihua
; APPLICANT: Cao, Yongwei
; APPLICANT: Wu, Wei
; APPLICANT: Boukharov, Andrey A.
; APPLICANT: Barbazuk, Brad
; APPLICANT: Li, Ping
; TITLE OF INVENTION: Rice Nucleic Acid Molecules and Other Molecules Associated With
; TITLE OF INVENTION: Plants and Uses Thereof for Plant Improvement
; FILE REFERENCE: 38-21(53221)B
; CURRENT APPLICATION NUMBER: US/10/437,963
; CURRENT FILING DATE: 2003-05-14
; NUMBER OF SEQ ID NOS: 204966
; SEQ ID NO 136993
; LENGTH: 372
; TYPE: PRT
```



```
; ORGANISM: Oryza sativa
; FEATURE:
; OTHER INFORMATION: Clone ID: PAT_MRT4530_38519C.1.pep
US-10-437-963-136993

Query Match      28.0%; Score 492; DB 16; Length 372;
Best Local Similarity 30.7%; Pred. No. 6.8e-40;
Matches 102; Conservative 81; Mismatches 115; Indels 34; Gaps 5;

QY 3 KALVIINCIILAIIGNCGGPLIMRLYFNNGGKRIWFSTFLETAGFPVIFIPLL----- 54
Db 19 RLLLFANYAALLGGSVASSLSRYVFAHGCHNRWVATLQVGVFPFLVPVYAGRSPSQP 78
QY 55 --PSYITRRSRNNVGDSSTFELIKPRLLIAAVIVGILSGFDNYLYAYGIAYLPVSTAALI 112
Db 79 RPPAFWFR-----RLLAACVVIGVLMGVNLLFSYSSSYPVSTSSL 121
QY 113 IASQLAFIAIFSPFMVKHKFTPTTINAVLLTVGAAVLGMHTET--DKPVHETHKQYITG 170
Db 122 LSTQLAFTLVLAIIVRHPLNFSNLNAVLLTSSVLIARSDSGEQGGSRARYFVG 181
QY 171 FLITVAANVMYAFILPLVELAYQAKQTMSTYTLVLEFQLILCLLASIVSVIGMFIAGDFK 230
Db 182 FAVTLGAAGLFAAYLPWMLVYRKA-VSGGFRNAVEQVIMQAAANTALAVAGLAAGGWK 240
QY 231 QALPKEARBEFKLGEALFYVAVFSATIMQGFILGAIGLIFSTSLVSGIMISVLLPITEV 290
Db 241 EELAR----WDLSPAAYWAVLAALVATWQACLMGTAGMVTLSLHSGVCMTAVLTANVI 296
QY 291 LAVIFYHEKFAEKGLSLALSLMGFVSFYGE 322
Db 297 GGVVVFRDPFGADKAVATVLCWNGFSSYLYGE 328
```

Search completed: September 24, 2004, 00:06:42  
Job time : 86.017 secs

**THIS PAGE BLANK (USPTO)**

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 Compugen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:40:51 ; Search time 14.9153 Seconds  
(without alignments)  
1228.856 Million cell updates/sec

Title: US-09-913-767-9

Perfect score: 1760

Sequence: 1 MKVALVINCIILAINCGG.....QQEESQETQSSLSRPISec 352

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 141681 seqs, 52070155 residues

Total number of hits satisfying chosen parameters: 141681

Minimum DB seq length: 0

Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%

Maximum Match 100%

Listing first 45 summaries

Database : SwissProt\_42.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

#### SUMMARIES

Result No.	Score	Query Match %	Length	DB ID	Description
1	134	7.6	518	1 SP5B_BACSU	Q00758 bacillus su
2	130.5	7.4	311	1 YCAL_CLOKJ	P38943 clostridium
3	125.5	7.1	556	1 NU2M_PODAN	P15578 podospira a
4	121.5	6.9	343	1 Y841_METTH	O26929 methanobact
5	120	6.8	374	1 YDB1_SCHPO	Q10354 schizosacch
6	119.5	6.8	450	1 GNUT_PSEAP	Q92411 pseudomonas
7	118.5	6.7	369	1 YUEF_BACSU	O32035 bacillus su
8	115	6.5	312	1 YCXC_BACSU	Q08794 bacillus su
9	115	6.5	654	1 PSTA_MYCGE	P47651 mycoplasma
10	114	6.5	304	1 YJMB_BACSU	O30646 salmonella
11	114	6.5	459	1 YJMB_BACSU	O34961 bacillus su
12	112.5	6.4	330	1 YETK_BACSU	O31540 bacillus su
13	111.5	6.3	349	1 YB64_SYNY3	P73771 synecocyst
14	111.5	6.3	735	1 DUR3_YEAST	P33413 saccharomyc
15	111.5	6.3	820	1 NAHI_RAT	O26431 rattus norv
16	111	6.3	452	1 NORM_BACSU	O31855 bacillus su
17	110.5	6.3	445	1 YPNP_BACSU	P54181 bacillus su
18	110	6.2	353	1 PERM_ECOLI	P77406 escherichia
19	110	6.2	384	1 FCTI_HUMAN	Q96a29 homo sapien
20	110	6.2	400	1 TYRP_HAEMIN	P26437 haemophilus
21	109.5	6.2	820	1 NAHI_MOUSE	O61165 mus musculu
22	109	6.2	299	1 YIGM_ECOLI	P75515 escherichia
23	109	6.2	517	1 Y120_MYCPN	P75515 mycoplasma
24	108.5	6.2	393	1 NUPC_BACSU	P39141 bacillus su
25	108	6.1	282	1 Y125_THEMA	Q9wx99 thermotoga
26	108	6.1	350	1 YDBI_BACSU	P66004 bacillus su
27	108	6.1	581	1 YH81_SCHPO	Q9p5n4 schizosacch
28	107.5	6.1	451	1 NTPJ_ENTHR	P43440 enterococcu
29	107.5	6.1	494	1 NUON_BUCBP	O89a44 buchnera ap
30	107.5	6.1	534	1 COX1_RICBP	O54069 rickettsia
31	107.5	6.1	630	1 YD13_SCHPO	Q92341 schizosacch
32	107	6.1	532	1 YC18_HAEMIN	O57251 haemophilus
33	107	6.1	651	1 PSTA_MYCPN	P75185 mycoplasma

RESULT 1  
SP5B\_BACSU STANDARD; PRT; 518 AA.  
AC Q00758;  
DT 01-APR-1993 (Rel. 25, Created)  
DT 01-APR-1993 (Rel. 25, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Stage V sporulation protein B.  
GN SPOVB OR SPOIIF OR BSU27670.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RX MEDLINE=92078103; PubMed=1744050;  
RA Popham D.L., Stragier P.;  
RT "Cloning, characterization, and expression of the spoVB gene of  
RT Bacillus subtilis.";  
RL J. Bacteriol. 173:7942-7949(1991).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RA Tosato V., Bolotin A., Bertani I., Valentino I., Bruschi C.V.;  
RT "A 17.8 kb segment in the spoVB-nadC region of the Bacillus subtilis  
RT 168 chromosome: sequencing and ruv operon identification.";  
RN Submitted (DEC-1997) to the EMBL/GenBank/DBJ databases.  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borriass R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,  
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
RA Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,  
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,  
RA Fritz C.Y., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,  
RA Ghim S.Y., Glaeser P., Goffeau A., Guittey E.J., Grandi G.,  
RA Guiseppe G., Guy B.J., Haga K., Haelech J., Harwood C.R., Henaut A.,  
RA Hilbert H., Holsappel S., Hosono S., Hullo M.P., Itaya M., Jones L.,  
RA Joris B., Karamata D., Kasahara Y., Klaere-Blanchard M., Klein C.,  
RA Kobayashi Y., Koester P., Koningsstein G., Krogh S., Kumano M.,  
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
RA Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
RA Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
RA Pressac E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
RA Sorokin A., Taccioni E., Takagi T., Takahashi H., Takemaru K.,  
RA Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzenegger T.,

#### ALIGNMENTS

P07869 bacillus su  
Q05394 cryptosporus  
Q99191 salmonella  
P94488 bacillus su  
Q01345 oncorhynchus  
P48761 cricetus  
Q56072 salmonella  
P42394 buchnera ap  
Q01650 homo sapien  
Q92127 mus musculu  
P33031 escherichia  
Q9ux84 sulfolobus

34 106.5 6.1 365 1 GRAB\_BACSU  
35 105.5 6.0 352 1 FMLR\_RABIT  
36 105.5 6.0 432 1 RFBX\_SALTI  
37 105.5 6.0 463 1 YNAJ\_BACSU  
38 105.5 6.0 759 1 NAHB\_ONCMY  
39 105.5 6.0 822 1 NAHI\_CRIGR  
40 105 6.0 299 1 YAMA\_BUCAP  
41 105 6.0 300 1 Y270\_BUCAP  
42 105 6.0 507 1 LATI\_HUMAN  
43 105 6.0 512 1 LATI\_MOUSE  
44 104.5 5.9 400 1 NUPC\_ECOLI  
45 104.5 5.9 469 1 SECY\_SULSO

RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
RA Yoshida K., Yoshikawa H.F., Zumbstein E., Yoshikawa H., Danchin A.;  
RT "The complete genome sequence of the Gram-positive bacterium Bacillus  
RT subtilis.",  
RL Nature 390:249-256(1997).  
CC -!- FUNCTION: SPOB PROTEIN IS INVOLVED, DIRECTLY OR INDIRECTLY, IN  
CC SPORE CORTEX BIOSYNTHESIS AND AFFECTS ONLY INDIRECTLY EXPRESSION  
CC OF LATE SPOREULATION GENES.  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -!- DEVELOPMENTAL STAGE: SPOB TRANSCRIPTION TAKES PLACE DURING THE  
CC SECOND HOUR OF SPOREULATION. IT MAY BE TRANSCRIBED MAINLY, IF NOT  
CC ONLY, IN THE MOTHER CELL. INDEED, IT IS REQUIRED ONLY IN THE  
CC MOTHER CELL.  
CC -!- SIMILARITY: Belongs to the polysaccharide synthase family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; M76666; AAA22806.1; -;  
CC EMBL; Y15896; CAB75337.1; -;  
CC EMBL; Z99118; CAB14727.1; -;  
CC PIR; I39989; I39989.  
CC Subtilisin; BG10778; spoVB.  
CC InterPro; IPR002797; Polysacc\_synt.  
CC Pfam; PF01943; Polysacc\_synt; 1.  
CC Sporeulation; Transmembrane; Complete proteome.  
FT TRANSMEM 6 26 POTENTIAL.  
FT TRANSMEM 45 65 POTENTIAL.  
FT TRANSMEM 91 111 POTENTIAL.  
FT TRANSMEM 120 140 POTENTIAL.  
FT TRANSMEM 165 185 POTENTIAL.  
FT TRANSMEM 186 206 POTENTIAL.  
FT TRANSMEM 250 270 POTENTIAL.  
FT TRANSMEM 281 301 POTENTIAL.  
FT TRANSMEM 326 346 POTENTIAL.  
FT TRANSMEM 348 368 POTENTIAL.  
FT TRANSMEM 387 407 POTENTIAL.  
FT TRANSMEM 411 431 POTENTIAL.  
FT TRANSMEM 446 466 POTENTIAL.  
FT TRANSMEM 478 498 POTENTIAL.  
SQ SEQUENCE 518 AA; 56078 MW; 2AE0A0F73FC4BE2B CRC64;  
  
Query Match 7.6%; Score 134; DB 1; Length 518;  
Best Local Similarity 21.7%; Pred. No. 0.024; Indels 117; Gaps 17;  
Matches 85; Conservative 53; Mismatches 137;  
  
QY 33 KRWFSTFLETAGFPVIFPLPSYITRRSRNNGDSTSF-----LIKPLLLIAAVIV 86  
DB 82 KNILVMSLITGVLSLIFPLFPAPVMAETLTDKRTLYPLLAITPVVPIAISSVLR 141  
QY 87 GILSGFDN-----YLAYGTAYLPVSTAAIIIASQLAFIAI 122  
DB 142 GYFQGNMPLANSQVLEQVRSILVAVCTTIFLPYIGIEYAAAGAMLSSVAGELASL-L 200  
QY 123 FSFFMVKHFT-----PFT-INAVLLTVGAVLGMHTETDKPVHETH- 164  
DB 201 YLEVCPRKTKIKRKHFTQSKNGKQFTTQLMSVSLPTGSRFIGNLSWFFPIVVAOS 260  
QY 165 -----KOY-ITGF-----LITVAAMVAFILPLV-----ELAYOKAKQWSTYL 203  
DB 261 LAIAGVATVAATKQYBELTGAFMTLLTLPSTFITYSLTALVPAISSEMEQKKLQVVEYR- 319  
QY 204 VLEFQLICLLASIVSGIWMFIAGDFKQALPKAREFKLGEA-----LFYV 249  
DB 320 -LEQAMRLCLLSGIVSVILFVFD-----ELMRVYGSSGAIVFKMAPFFLLYTF 371  
QY 250 VAVFSAII-----WQGFLLGAI-----GLIFSTSLVS-GIMISVLLPITEVLAVIFY 296

DB 372 QGFLQAVLQALNLAGAMNLSLIGALVTKTLIFVLATRPSLGMGAALAIVTGMVLVTL 431  
QY 297 HEKFOA-----EKGLSLA-LSLWGFVS 317  
DB 432 HAATSVKVLPSIKIKIKEYALSPFAVIVICGFIS 463  
  
RESULT 2  
YCAP CLOKL STANDARD; PRT; 311 AA.  
AC P38943;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 28-FEB-2003 (Rel. 41, Last annotation update)  
DE Hypothetical transport protein in catl 5' region (ORF).  
OS Clostridium kluyveri.  
OC Bacteria; Firmicutes; Clostridia; Clostridiales; Clostridiaceae;  
OC Clostridium.  
OX NCBI\_TaxID=1534;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=DSM 555;  
RC MEDLINE=96146540; PubMed=8550525;  
RA Soehling B., Gottschalk G.;  
RT "Molecular analysis of the anaerobic succinate degradation pathway in  
RT Clostridium kluyveri".  
RL J. Bacteriol. 178:871-880(1996).  
CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -!- SIMILARITY: Belongs to the eamA transporter family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
CC  
CC EMBL; L21902; AAA92345.1; -;  
CC InterPro; IPR000620; DUF6.  
CC Pfam; PF00892; DUF6; 2.  
CC Hypothetical protein; Transport; Transmembrane.  
FT TRANSMEM 6 26 POTENTIAL.  
FT TRANSMEM 33 53 POTENTIAL.  
FT TRANSMEM 70 90 POTENTIAL.  
FT TRANSMEM 97 117 POTENTIAL.  
FT TRANSMEM 123 143 POTENTIAL.  
FT TRANSMEM 155 175 POTENTIAL.  
FT TRANSMEM 185 205 POTENTIAL.  
FT TRANSMEM 219 239 POTENTIAL.  
FT TRANSMEM 244 264 POTENTIAL.  
FT TRANSMEM 265 285 POTENTIAL.  
SQ SEQUENCE 311 AA; 33507 MW; A5E7CB9EC9ADF0B CRC64;  
  
Query Match 7.4%; Score 130.5; DB 1; Length 311;  
Best Local Similarity 16.7%; Pred. No. 0.026;  
Matches 53; Conservative 62; Mismatches 126; Indels 77; Gaps 10;  
  
QY 48 VIFPLPSYITRRSRNNGDSTSFLLIKPRLIIAAVIVGILS-GFDNLYAYGTAYLPV 106  
DB 46 ILLPMVAKDKIKVKLKLTDGNDLGYL-----ALCGILAVSISMSMLQFATYTKA 95  
QY 107 STAAIIIASQLAFTAFISFFFMVKHFTPTTINAVLLTVGAALV-----GMHTETDK 158  
DB 96 STAAVLCTNAVFTTPFAYFILKEIKIGITIVSIIVSLIGVVIIPNPAKVMIEGIGSRD- 154  
QY 159 PVHETHKQYITGFLITVAAAVVAFILPLVELAYOKAKQWSTYLVLE-----FQ 208  
DB 155 -----LIGICPALVAADVWSL-----YTVISKRIEYGYGVFN 188  
QY 209 LILCLLASIVSGIWMFIAGDFKQALPKAREFKLGEALFYVAVFSAIIWQGFLLGAI 268  
DB 189 CISFFFGVALLILLVTVG-----RPIFSGITLNNIL-----VLLYMGIFIKAVGY 234

269 IF-----STSLVSGMISVLPITETVLAVIFHYEKFQAEKGLSLALSGLWGFVSFYGE 322  
 235 ICYLGAKETSASVTASTVFLIKPALATVLAAILGESIEVNVIGVFIIGSI-INYS 293  
 323 IKSGEDKRIQESQET 340  
 294 NKKANDLKRVTSSAES 311

RESULT 3  
 NU2M PODAN STANDARD; PRT; 556 AA.  
 AC F15578;  
 DT 01-APR-1990 (Rel. 14, Created)  
 DT 01-APR-1990 (Rel. 14, Last sequence update)  
 DT 15-JUL-1999 (Rel. 38, Last annotation update)  
 DE NADH-ubiquinone oxidoreductase chain 2 (EC.1.6.5.3).  
 GN ND2.  
 OS Podospora anserina.  
 OC Mitochondrion.  
 OC Eukaryota; Fungi; Ascomycota; Pezizomycotina; Sordariomycetes;  
 OC Sordariomycetidae; Sordariales; Lasiosphaeriaceae; Podospora.  
 OX NCBI\_TaxID=5145;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=s, and A;  
 RX MEDLINE=89125610; PubMed=2975708;  
 RA Cummings D.J., Domenico J.M.;  
 RT "Sequence analysis of mitochondrial DNA from Podospora anserina.  
 RT Pervasiveness of a class I intron in three separate genes.";  
 RL J. Mol. Biol. 204:815-839 (1998).  
 RN [2]  
 RP COMPLETE GENOME.  
 RC STRAIN=s;  
 RX MEDLINE=90291512; PubMed=2357736;  
 RA Cummings D.J., McNally K.L., Domenico J.M., Matsuura E.T.;  
 RT "The complete DNA sequence of the mitochondrial genome of Podospora  
 RT anserina.";  
 RL Curr. Genet. 17:375-402 (1990).  
 CC -1- CATALYTIC ACTIVITY: NADH + ubiquinone = NAD(+) + ubiquinol.  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Mitochondrial  
 CC inner membrane.  
 CC -1- SIMILARITY: Belongs to the complex I subunit 2 family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; X55026; CAA38765.1; -;  
 CC EMBL; X14485; CAA32646.1; -;  
 CC F1R; S02154; S02154.  
 CC InterPro; IPR001750; Oxidored g1.  
 CC Pfam; PF00361; oxidored g1; 1.  
 KW Oxidoreductase; NAD; Ubiquinone; Mitochondrion; Transmembrane.  
 SQ SEQUENCE 556 AA; 62554 MW; 5613ACB0D2324BAF CRC64;

Query Match 7.1%; Score 125.5; DB 1; Length 556;  
 Best Local Similarity 24.7%; Pred. No. 0.099;  
 Matches 69; Conservative 51; Mismatches 110; Indels 49; Gaps 13;

38 STPLETAGPVPVPIPLPILLSVITRRNNVGDSTSPFLIKERLLIA---AVIVGILSGFDN 94  
 273 TTFVAIIAKISIFILL--ELVYHTNYLSEFSTYTL-----LLISLFSLIIGTVVGLTQ 326

95 YLYVAYGIAPVPTAALI-----IASQLAIFSPFPMVKHKPTPTINAVLLTWG 146  
 327 FRIKRLAYSTISHVGFILLALSGCSIESIQAFI-----FYLIQYSISNLNV-FIIITIG 381

147 AAVLGMHT-----ETDKPVHETHKQ-----YITGFLITVAAAVMYAF--ILPLVEL 130  
 382 FSLGYITTNKEYKDLKNNSPQIVISQLKGYFIYNPLLSLSLAIITFSFVGIPPLVGF 441  
 191 AYOKAKOTMSYTLVLEFQILCLLASIVSVIG-MFIAGDFKQA---LPKREARFKLGEAL 246  
 442 F---AKOMVLSAALDNGYIFLTIAITLTSVIGAVYVNIKKIFFYLPDHSINPSIGEFL 498  
 247 FYVAVFSALIMQGFELGALIFSTSSLSVSGMISVLL 285  
 499 FKGLIFEA---GDFKGRITLSSPFSITISITLIVL 533

RESULT 4  
 Y841 METTH STANDARD; PRT; 343 AA.  
 AC O26929;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical transport protein MTH841.  
 GN MTH841.  
 OS Methanobacterium thermoautotrophicum.  
 OC Archaea; Euryarchaeota; Methanobacteria; Methanobacteriales;  
 OC Methanobacteriaceae; Methanothermobacter.  
 OX NCBI\_TaxID=187420;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=Delta H;  
 RX MEDLINE=98037514; PubMed=9371463;  
 RA Smith D.R., Doucette-Stamm L.A., Deloughery C., Lee H.-M., Dubois J.,  
 RA Aldredge T., Bashirzadeh R., Blakely D., Cook R., Gilbert K.,  
 RA Harrison D., Hoang L., Keagle P., Lumm W., Pothier B., Qiu D.,  
 RA Spadafora R., Vicare R., Wang Y., Wierzbowski J., Gibson R.,  
 RA Jiwan N., Caruso A., Bush D., Safer H., Patwell D., Prabhakar S.,  
 RA McDougall S., Shimer G., Goyal A., Pietrowski S., Church G.M.,  
 RA Daniels C.J., Mao J.-I., Rice P., Noelling J., Reeve J.N.;  
 RT "Complete genome sequence of Methanobacterium thermoautotrophicum  
 RT deltaH: functional analysis and comparative genomics.";  
 RL J. Bacteriol. 179:7135-7155 (1997).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -1- SIMILARITY: Belongs to the eamA transporter family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC  
 CC EMBL; AB000861; BAB85339.1; -;  
 CC F1R; C69212; C69212.  
 CC InterPro; IPR000620; DUF6.  
 CC Pfam; PF00892; DUF6; 2.  
 KW Hypothetical protein; Transport; Transmembrane; Complete proteome.  
 FT TRANSMEM 13 33 POTENTIAL.  
 FT TRANSMEM 44 64 POTENTIAL.  
 FT TRANSMEM 71 91 POTENTIAL.  
 FT TRANSMEM 121 141 POTENTIAL.  
 FT TRANSMEM 148 168 POTENTIAL.  
 FT TRANSMEM 177 197 POTENTIAL.  
 FT TRANSMEM 203 223 POTENTIAL.  
 FT TRANSMEM 244 264 POTENTIAL.  
 FT TRANSMEM 269 289 POTENTIAL.  
 FT TRANSMEM 296 316 POTENTIAL.  
 FT TRANSMEM 320 340 POTENTIAL.  
 SQ SEQUENCE 343 AA; 37839 MW; 4296C7890CB91B7F CRC64;

Query Match 6.9%; Score 121.5; DB 1; Length 343;  
 Best Local Similarity 23.1%; Pred. No. 0.12;  
 Matches 68; Conservative 50; Mismatches 126; Indels 51; Gaps 12;

QY 44 AGFPVIFI---PL---LFSYITRRSNNVGDSTGFFLKPRLIAAIVVIGLSDNLY 97  
 DB 84 AGIFLFFIRETPLEHVLNAINRKGESARIKRDYLI---LTLTALLSTVIAPL---LF 137  
 QY 98 AYGTAIYLPVSTAAIIASQALAFIAFSPFMVKKHFTPTTINAVLLTVGAIVLGMHTED 157  
 DB 138 LTGLGDTTAVNASLTLNVEVFIILLGLYLIPTETLQKDFIGVLIILGAVYL---LIEGD 195  
 QY 158 KPVHETHKQYITGFLITVAANAAMVAFILPLVELAYQKAKQMTSYTLVLEFOLILCLLASI 217  
 DB 196 PST-ILRNVAVTGNFLVNAAPFNSLDTVLSKFLSRKED-----LIFISGV 240  
 QY 218 VSVIG-----MFTAGDFKQALPKAREFKLGEALFVVAVFSAIIVWQGFLLCA--IGL 268  
 DB 241 KSSVGGFVLLIIMJLG-INTLEPLEMLPYALGVSVFSIGCSFTLIYIAIREIGASVMGA 299  
 QY 269 IFTSTSLVSGMISVLLPITEVLAVIFVHEKFOAEKGLSLALSIMG-FVSYFYGE 322  
 DB 300 LFFLSLFGAI-----FAPILREPSIMQGISGIVMLTGVFILIWNCK 343

RESULT 5

YDB1 SCHPO STANDARD; PRT; 374 AA.  
 AC Q10354; Q9URX5;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 16-OCT-2001 (Rel. 40, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Hypothetical protein C22E12.01 in chromosome 1.  
 GN SPAC22E12.01 OR SPAC890.09  
 OS Schizosaccharomyces pombe (Fission yeast).  
 OC Eukaryota; Fungi; Ascomycota; Schizosaccharomycetes;  
 OC Schizosaccharomycetales; Schizosaccharomycetaceae;  
 OC Schizosaccharomycetes.  
 OX NCBI\_TaxID=4896;  
 RN [1]  
 RP SEQUENCE FROM N.A.

RC STRAIN=972;  
 RX MEDLINE=21848401; PubMed=11859360;  
 RA Wood V., Williams R., Rajandream M.A., Lyne M., Lyne R., Stewart A.,  
 RA Sgouras J., Peat N., Hayles J., Baker S., Basham D., Bowman S.,  
 RA Brooks K., Brown D., Brown S., Chillingworth T., Church C.M.,  
 RA Collins M., Connor R., Cronin A., Davis P., Feltwell T., Fraser A.,  
 RA Gentles S., Goble A., Hamlin N., Harris D., Hidalgo J., Hodgson G.,  
 RA Holroyd S., Hornsby T., Howarth S., Huckle E.J., Hunt S., Jagels K.,  
 RA James K., Jones L., Jones M., Leather S., McDonald S., McLean J.,  
 RA Mooney P., Moule S., Mungall K., Murphy L., Niblett D., Odell C.,  
 RA Oliver K., O'Neil S., Pearson D., Quail M.A., Rabinowitsch E.,  
 RA Rutherford K., Rutter S., Saunders D., Seeger K., Sharp S.,  
 RA Skelton J., Simmonds M., Squares R., Squares S., Warren T., Whitehead S.,  
 RA Taylor K., Taylor R.G., Tivey A., Walsh S.V., Warren T., Whitehead S.,  
 RA Woodward J., Voiclaert G., Aert R., Robben J., Grymonprez B.,  
 RA Weltjens I., Vanstreels E., Rieger M., Schaefer M., Mueller-Auer S.,  
 RA Gabel C., Fuchs M., Fritzc C., Holzer E., Moestl D., Hilbert H.,  
 RA Borzym K., Langer I., Beck A., Lehar H., Reinhardt R., Pohl T.M.,  
 RA Eger P., Zimmermann W., Wedler H., Wambutt R., Purnelle B.,  
 RA Goffeau A., Cadieu E., Dreano S., Gloux S., Lelaure V., Mottier S.,  
 RA Galibert F., Aves S.J., Xiang Z., Hunt C., Moore K., Hurst S.M.,  
 RA Lucas M., Rochet M., Gaillardin C., Tallada V.A., Garzon A., Thode G.,  
 RA Daga R.R., Cruzado L., Jimenez J., Sanchez M., del Rey F., Benito J.,  
 RA Dominguez A., Revuelta J.L., Moreno S., Armstrong J., Forsburg S.L.,  
 RA Carrutti L., Lowe T., McCombie W.R., Paulsen I., Potashkin J.,  
 RA Shipakowski G.V., Ussery D., Barrell B.G., Nurse P.;  
 RA Nature 415:871-880 (2002).  
 RL The genome sequence of *Schizosaccharomyces pombe*."

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announcement/>)

CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch).  
 CC EMBL; AL133498; CAB63500.1; --  
 DR EMBL; Z70043; CAA93888.1; --  
 DR PIR; T50265; T50265.  
 DR GeneDB SPombe; SPAC22E12.01; --  
 DR GO; GO:0016020; C:membrane; ISS.  
 DR GO; GO:0006810; P:transport; ISS.  
 DR KW Hypothetical protein; Transmembrane.  
 FT TRANSMEM 50 70 POTENTIAL.  
 FT TRANSMEM 151 171 POTENTIAL.  
 FT TRANSMEM 175 195 POTENTIAL.  
 FT TRANSMEM 201 221 POTENTIAL.  
 FT TRANSMEM 238 258 POTENTIAL.  
 FT TRANSMEM 273 293 POTENTIAL.  
 FT TRANSMEM 329 349 POTENTIAL.  
 SQ SEQUENCE 374 AA; 42300 MW; AE504D822FAF8B12 CRC64;

Query Match 6.8%; Score 120; DB 1; Length 374;  
 Best Local Similarity 23.1%; Pred. No. 0.16;  
 Matches 82; Conservative 60; Mismatches 163; Indels 50; Gaps 15;

QY 2 VKALVIINCIILA--IGNCGPLIMLYFNNGKRIWFTFTETAGFPVIFILLESYIT 59  
 DB 47 ITRVITVILVLAIFYFSLLSMMNKWIFSEKMDQFPLFLSSC---QMLVOMGFALIT 103  
 QY 60 ----RRSNVGDSTSPFLIKPRLLIAAVIVIGILSGFDNLYAYGIAVLPVSTAALIAS 115  
 DB 104 ILAPPRYPQNKKNFSLWLEYFYRAGICALVIGLDIGLSN----ASLETITLSFVTMCRSS 159  
 QY 116 QLAFIAIFSPFMVKKHFTPTTINAVLLTVGAIVLGMHTETDKPVHETHKQYI-TGLFIT 174  
 DB 160 ILIFVFFSVIFRIEMF-DWILLCITLIVISAGVVLVMTATET-----QFVLSGFLV 209  
 QY 175 VAAAVMYAFILPLVE---LAYQKAKQMTSYTLVLEFOLILCLLASIVSVIGMFIAGDFKQ 231  
 DB 210 MASSVLSGLRWALTQKLLDHPMTSNPFTSLFALPTLMFLFL---VAGLIFEGVPVRF 264  
 QY 232 ALPKAREEKLGEALFYVAVFSAIIVWQGFLLCA--IGLIFSTSSL---VSGIMISVLLP 286  
 DB 265 IESPAWKEF--GPMSVVLVPGTL---AFFWASEFGLIKTSIVTLVSGILKEI--- 316  
 QY 287 ITEVLAVIFVHEKFOAEKGLSLALSIMGFVSFYFGEIKSGEDKKRIQOESQETE 341  
 DB 317 ITLITASTLFYHDILLINIVGLVITLQGVYNYRITKGNK-----EAEKEYE 366

RESULT 6

GNUT\_PSEAE STANDARD; PRT; 450 AA.  
 AC Q9Z1J1;  
 DT 30-MAY-2000 (Rel. 39, Created)  
 DT 30-MAY-2000 (Rel. 39, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Gluconate permease.  
 GN GNUT OR PA2322.  
 OS Pseudomonas aeruginosa.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Pseudomonadales;  
 OC Pseudomonadaceae; Pseudomonas.  
 OX NCBI\_TaxID=287;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RA Hager P.W., Covert-Rinaldi A., Wallace W.H., Phipps P.V. Jr.;  
 RT "Cloning and sequence analysis of the gluconate operon of *Pseudomonas*  
 RT *aeruginosa* PAO.";  
 RL Submitted (SEP-1997) to the EMBL/GenBank/DBJ databases.  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 15692 / PA01;  
 RX MEDLINE=20437337; PubMed=10984043;  
 RA Stover C.K., Pham X.-Q.T., Erwin A.L., Mizoguchi S.D., Warren P.,  
 RA Hickey M.J., Brinkman F.S.L., Hufnagle W.O., Kowalik D.J., Lagrou M.,

Garber R.L., Goltry L., Tolentino E., Westbrook-Wadman S., Yuan Y., Brody L.L., Coulter S.N., Folger K.R., Kas A., Larbig K., Lim R.M., Smith K.A., Spencer D.H., Wong G.K.-S., Wu Z., Paulsen I.T., Reizer J., Sailer M.H., Hancock R.E.W., Lory S., Olson M.V., "Complete genome sequence of *Pseudomonas aeruginosa* PAO1, an opportunistic pathogen.", Nature 406:959-964(2000).

RT Nature 406:959-964(2000).

CC -1- PATHWAY: Glucuronate utilization.

CC -1- SUBCELLULAR LOCATION: Integral membrane protein. Inner membrane (Potential).

CC -1- SIMILARITY: BELONGS TO THE GNP FAMILY OF PERMEASES.

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL; AF026470; AAD01803.1; -

DR EMBL; AE004658; AAG05710.1; -

DR PIR; E83355; E83355.

DR InterPro; IPR003474; GlnC transporter.

DR Pfam; PF02447; GntP\_permease; 1.

DR TIGRFAMs; TIGR00791; gntP; 1.

KW Glucuronate utilization; Sugar transport; Transport; Transmembrane; Inner membrane; Complete proteome.

KW TRANSMEM 6 26 POTENTIAL.

FT TRANSMEM 30 50 POTENTIAL.

FT TRANSMEM 60 80 POTENTIAL.

FT TRANSMEM 116 136 POTENTIAL.

FT TRANSMEM 142 162 POTENTIAL.

FT TRANSMEM 183 203 POTENTIAL.

FT TRANSMEM 233 253 POTENTIAL.

FT TRANSMEM 269 289 POTENTIAL.

FT TRANSMEM 312 332 POTENTIAL.

FT TRANSMEM 338 358 POTENTIAL.

FT TRANSMEM 366 386 POTENTIAL.

FT TRANSMEM 430 450 POTENTIAL.

SQ SEQUENCE 450 AA; 47296 MW; AED12865DBD97436 CRC64;

Query Match 6.8%; Score 119.5; DB 1; Length 450;

Best Local Similarity 21.8%; Pred. No. 0.21;

Matches 87; Conservative 52; Mismatches 139; Indels 121; Gaps 15;

QY 11 ILLAIGCGPL-----IMRLFNNGKKEIWSTPLET--AGFPVIF----- 50

DB 71 VILALGMLGLKWDGSGAQIARTLRP---GKRVHWSMMLAFLVGLPLFFETGFI 127

QY 51 --IPLFSYITRRSNVGDSTSPFLIK---PRLLIAAVTIGILSGPDNLYAYGIAYLP 105

DB 128 LLPLVP--IVARRS-----GVSLIKIGIPLLAGLSAVHGLVPPHPPGFLAIGVPGAD 178

QY 106 VSTA---ALIIASQAFIAIFSPFMVKHPTPTTINAVVLTGCAAVLGMHTTDPKVE 162

DB 179 IGKTLVGLVLPALPTAIAIGLFCALVSRTGTPPSAELV-----EQIAHE 224

QY 163 THQYITGFLITVAAYVAFILPL-----VELAYOKA 195

DB 225 PETQDLPSFGVTLATVLLPFLMLLKTTFADVPDGHAFRAMDMIGHIPISALLALLVA 284

QY 196 KQWSTYTLVLEFOLICLL-----ASIVSVIGMFIAGDFKQALPKAREEKLGE----- 244

DB 285 LYTFGVARGDFSKILLRLDQSLAPTAIWNIG--AGGFKQMLVASGVGDVIGHUAVN 342

QY 245 -----ALFYVAVFSAIIWQGFPLGAI---GLIFSTSLVSGIMISVLLPITEVLAVIF 295

DB 343 AQISPELLAMLVAAVIRIATGSAVTATITGAGIVVPVIDLIPGVNRELLVATGAGSLIL 402

QY 296 YH-----EKQAEKGLSLALSLSLQGV 316

DB 403 SHVNDAGFMLVKQYFNMSVSETEKTTWTAMETILSVGLV 441

## RESULT 7

YUEF\_BACSU STANDARD; PRT; 369 AA.

AC Q32035; 2001 (Rel. 40, Created)

DT 16-OCT-2001 (Rel. 40, Last sequence update)

DT 10-OCT-2003 (Rel. 42, Last annotation update)

DE Hypothetical protein yuef.

GN YUEF OR BSU31800.

OS Bacillus subtilis.

OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.

OX NCBI\_TaxID=1423; [1]

RN SEQUENCE FROM N.A.

RP STRAIN=168;

RX MEDLINE=98044033; PubMed=9384377;

RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G., Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S., Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S., Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M., Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A., Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T., Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D., Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N., Ghim S.Y., Glaser P., Goffeau A., Golightly E.J., Grandi G., Guiseppe G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A., Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L., Joris B., Karamata D., Kasahara Y., Kjaer-Blanchard M., Klein C., Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M., Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V., Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C., Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M., Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H., Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M., Presecan E., Pujic P., Purnelle B., Rapoport G., Rey M., Reynolds S., Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y., Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F., Sakiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B., Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K., Takeuchi M., Tamakoshi A., Tanaka T., Terpstra P., Tognoni A., Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A., Viari A., Wambutt R., Wedler E., Wedler H., Weitzensegger T., Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K., Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.

RT "The complete genome sequence of the Gram-positive bacterium *Bacillus subtilis*."

RL Nature 390:249-256(1997).

CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).

CC -1- SIMILARITY: BELONGS TO THE UPF0118 (PERM) FAMILY.

CC -----

CC This SWISS-PROT entry is copyright. It is produced through a collaboration between the Swiss Institute of Bioinformatics and the EMBL outstation - the European Bioinformatics Institute. There are no restrictions on its use by non-profit institutions as long as its content is in no way modified and this statement is not removed. Usage by and for commercial entities requires a license agreement (See <http://www.isb-sib.ch/announce/> or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).

CC -----

EMBL; Z99120; CAB15168.1; -

DR PIR; G70007; G70007.

DR Subtilist; BG13961; Yuef.

DR InterPro; IPR02549; UPF0118.

DR Pfam; PF01594; UPF0118; 1.

KW Hypothetical protein; Transmembrane; Complete proteome.

FT TRANSMEM 13 33 POTENTIAL.

FT TRANSMEM 34 54 POTENTIAL.

FT TRANSMEM 73 93 POTENTIAL.

FT TRANSMEM 159 179 POTENTIAL.

FT TRANSMEM 213 233 POTENTIAL.

FT TRANSMEM 234 254 POTENTIAL.

FT TRANSMEM 271 291 POTENTIAL.

BEST AVAILABLE COPY

FT TRANSMEM 316 336 POTENTIAL.  
SQ SEQUENCE 369 AA; 40896 MW; 5F86C34CF258C4B8 CRC64;

Query Match  
Best Local Similarity 6.7%; Score 118.5; DB 1; Length 369;  
Matches 81; Conservative 58; Mismatches 116; Indels 123; Gaps 17;

QY 37 FSTFLETAGFPVIFPLH---FSYITRRSSNNVGDSTSFLLIKPRLLIAAVVIGLGF 93  
DB 33 FIVFISTLFFPMLIAGILVFIFPNVVRLEKKIPRTLSI-----LTIYLLFIGLL----- 82

QY 94 NYLYXGATYALPVSATLIIA-----SQLAFIAFSFFM----- 127  
DB 83 -----AFISAGVPIITAQVTGLFNPLPYIKQIQALTKDLSHSQWFTMMNQDVVS 134

QY 128 ---VKHKFPFTINAVLLTVG-AAVLGNHETDKPVHETHKQYITGLTIIVAAAVMYAF 183  
DB 135 ISKIQSLTSFLQNPQNTSLSAVGVVFN-----ITLVITVPFILFYML 182

QY 184 -----ILPLVELAYQ-----KAKQNTMSYTLVLEF--QLILCLLASIVSVIGMFI 225  
DB 183 KDGHRFPPLAVKILP---ASYRTEGLKIFKDLSDTLAAVYFQQLLICLVFGVTACFYGILI 239

QY 226 AGDFQALPKAREFKLGEALYVYVSAIIWQGFELCA-----IGLIFTSSILVSGIM 280  
DB 240 AG-----LP-----YALILGIVMAITNIIPYVGFELGAPAVIVGVFMDSPAKALFAI 287

QY 281 ISVLL-----PITEVLAVIFYHEKFOAEKGLSLALSIMGVFV-SYF 319  
DB 288 VVIVVQDGNLLSPLVIGKRLNTHPLTITILLI-GAGSFGGILGMILAVPVYAVVKAFF 346

QY 320 YGEIKSGEDKRIQOEE 337  
DB 347 LNVIRLIKLRQSRLEEN 364

RESULT 8  
YXCS BACSU STANDARD; PRT; 312 AA.  
AC Q08754;  
DT 01-FEB-1995 (Rel. 31, Created)  
DT 01-FEB-1995 (Rel. 31, Last sequence update)  
DT 10-OCT-2003 (Rel. 42, Last annotation update)  
DE Hypothetical transport protein yxcX.  
GN YXCX OR BSU03550.  
OS Bacillus subtilis.  
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.  
OX NCBI\_TaxID=1423;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168 / JH642;  
RX MEDLINE=93360813; PubMed=8355609;  
RA Cosmina P., Rodriguez F., de Ferza F., Grandi G., Perego M.,  
RA Vanena G., van Sinderen D.,  
RT "Sequence and analysis of the genetic locus responsible for surfactin  
RT synthesis in *Bacillus subtilis*."  
RL Mol. Microbiol. 8:821-831(1993).  
RN [2]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=97124189; PubMed=8969502;  
RA Yamane K., Kumano M., Kurita K.;  
RT "The 25 degrees-36 degrees region of the *Bacillus subtilis* chromosome:  
RT determination of the sequence of a 146 kb segment and identification  
RT of 113 genes."  
RL Microbiology 142:3047-3056(1996).  
RN [3]  
RP SEQUENCE FROM N.A.  
RC STRAIN=168;  
RX MEDLINE=98044033; PubMed=9384377;  
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,  
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,  
RA Borries R., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,

Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,  
Choi S.K., Codani J.J., Connerton I.F., Cummings N.J., Daniel R.A.,  
Entian K.D., Errington J., Fabret C., Ferrar E., Foulger D.,  
Fritz C., Fujita M., Fujita Y., Fuma S., Gallizzi A., Galleron N.,  
Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,  
Guiseppi G., Guy B.J., Haga K., Haiech J., Harwood C.R., Henaut A.,  
Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,  
Joris B., Karamant D., Kaehara Y., Klaerz-Blanchard M., Klein C.,  
Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,  
Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,  
Lee S.M., Levine A., Liu H., Masuda S., Mauel C., Medigue C.,  
Medina N., Mellado R.P., Mizuno M., Moestl D., Nakai S., Noback M.,  
Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,  
Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,  
Presecan E., Pujić P., Purnelle B., Rapoport G., Rey M., Reynolds S.,  
Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,  
Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,  
Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,  
Sorokin A., Tacconi E., Takagi T., Takahashi H., Takemaru K.,  
Takeuchi M., Tanakoshi A., Tanaka T., Terpstra P., Tognoni A.,  
Toato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,  
Viari A., Wambutt R., Wedler E., Wedler H., Weitzenecker T.,  
Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,  
Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,  
"The complete genome sequence of the Gram-positive bacterium *Bacillus  
subtilis*."  
Nature 390:249-256(1997).  
RL Nature 390:249-256(1997).  
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
CC -1- SIMILARITY: Belongs to the eamA transporter family.  
CC  
CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
CC between the Swiss Institute of Bioinformatics and the EMBL Outstation -  
CC the European Bioinformatics Institute. There are no restrictions on its  
CC use by non-profit institutions as long as its content is in no way  
CC modified and this statement is not removed. Usage by and for commercial  
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/  
CC or send an email to license@isb-sib.ch).  
CC  
CC EMBL; X70356; CAA49823.1; -;  
CC EMBL; D50453; BAA08989.1; -;  
CC EMBL; Z99105; CAB12149.1; -;  
CC PIR; I40491; I40491.  
CC Subtilist; BG10174; YXCX.  
CC InterPro; IPR000620; DUF6.  
CC Pfam; PF00892; DUF6; 2.  
CC Hypothetical protein; Transport; Transmembrane; Complete proteome.  
FT TRANSMEM 11 31  
FT TRANSMEM 46 66  
FT TRANSMEM 72 92  
FT TRANSMEM 98 118  
FT TRANSMEM 128 148  
FT TRANSMEM 155 171  
FT TRANSMEM 183 203  
FT TRANSMEM 221 241  
FT TRANSMEM 254 274  
FT TRANSMEM 277 297  
SQ SEQUENCE 312 AA; 34045 MW; FIC28D0374869C8C CRC64;

Query Match 6.5%; Score 115; DB 1; Length 312;  
Best Local Similarity 21.5%; Pred. No. 0.31;  
Matches 56; Conservative 43; Mismatches 78; Indels 84; Gaps 11;

QY 43 TAGFPVIFPLFSYITRR-RSSNVGDSTSFLLIKPRLLIAAVVIGLGFNDLYAYGI 101  
DB 45 TIAFAAAVPIPLFGWKLKSLVLDILPLALLYPALFFS-----PQAFGL 91

QY 102 AYLVPVSTAALIASQAFIAIFSPFMVKHPTPTINAVLLTVGAIVLGMHETDKPVH 161  
DB 92 VYSSSSSEAGIIQAAIPITFTMVFAAVLKER-----PTWTKQ--- 127

QY 162 ETHKQYITGFLITVAAAQVYAFILPLVLAQKAKOTMSYTLVLEFQILCLLASIVSVI 221



Db 128 -----GFTVLSVAGVMFIVNKGVDVESASLKGS-----LILILSALSSAM 168  
 QY 222 GMFIAGDFKQALPKAREFKLGEALFYVAVFSAIWOQFGLGALFISTSSIV-----S 277  
 Db 169 -----YNTAARQWQRFKLTETY-----IMSAI---GF-----VVFNAIALVRHGA 208  
 QY 278 GIMISVLLPITE---VLAVIP 295  
 Db 209 GTGTGYFLPFRPGFVLAIVY 229

RESULT 9  
 PSTA\_MYCGE  
 ID PSTA\_MYCGE STANDARD; PRT; 654 AA.  
 AC P47651;  
 DT 01-OCT-1996 (Rel. 34, Created)  
 DT 01-OCT-1996 (Rel. 34, Last sequence update)  
 DE 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE Phosphate transport system permease protein psta homolog.  
 GN PSTA OR MG411.  
 OS Mycoplasma genitalium.  
 OC Bacteria; Firmicutes; Mollicutes; Mycoplasmataceae; Mycoplasma.  
 OX NCBI\_TaxID=2097;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=96026346; PubMed=7569993;  
 RA Fraser C.M., Gocayne J.D., White O., Adams M.D., Clayton R.A.,  
 RA Fleischmann R.D., Bult C.J., Kerlavage A.R., Sutton G., Kelley J.M.,  
 RA Fritchman J.L., Weidman J.F., Small K.V., Sandusky M., Fuhmann J.L.,  
 RA Nguyen D.T., Uterback T.R., Saudek D.W., Phillips C.A., Merrick J.M.,  
 RA Tomb J.-F., Dougherty B.A., Bost K.F., Hu P.-C., Lucier T.S.,  
 RA Peterson S.N., Smith H.O., Hutchison C.A. III, Venter J.C.;  
 RT "The minimal gene complement of Mycoplasma genitalium.";  
 RL Science 270:397-403(1995).  
 RN [2]  
 RP SEQUENCE OF 136-211 FROM N.A.  
 RC STRAIN=ATCC 33530 / G-37;  
 RX MEDLINE=94075230; PubMed=8253680;  
 RA Peterson S.N., Hu P.-C., Bost K.F., Hutchison C.A. III;  
 RT "A survey of the Mycoplasma genitalium genome by using random  
 sequencing.";  
 RL J. Bacteriol. 175:7918-7930(1993).  
 CC -1- FUNCTION: COULD BE PART OF A BINDING-PROTEIN-DEPENDENT TRANSPORT  
 CC SYSTEM FOR PHOSPHATE; PROBABLY RESPONSIBLE FOR THE TRANSLOCATION  
 CC OF THE SUBSTRATE ACROSS THE MEMBRANE (BY SIMILARITY).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: Belongs to the binding-protein-dependent transport  
 CC system permease family. Cysw subfamily.  
 CC -----  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See <http://www.isb-sib.ch/announce/>  
 CC or send an email to [license@isb-sib.ch](mailto:license@isb-sib.ch)).  
 CC -----  
 DR EMBL; U39723; AAC71639.1; -;  
 DR EMBL; U01746; AAD10559.1; -;  
 DR PIR; E64245; E64245.  
 DR TIGR; MG411; -;  
 DR InterPro; IPR000515; BPD transp.  
 DR Pfam; PF00528; BPD transp; 2.  
 DR PROSITE; PS00928; ABC\_TM1; 2.  
 KW Transport; Phosphate transport; Transmembrane; Complete proteome.  
 FT TRANSMEM 22 42 POTENTIAL.  
 FT TRANSMEM 64 84 POTENTIAL.  
 FT TRANSMEM 113 133 POTENTIAL.  
 FT TRANSMEM 143 163 POTENTIAL.  
 FT TRANSMEM 368 388 POTENTIAL.  
 FT TRANSMEM 418 438 POTENTIAL.  
 FT TRANSMEM 454 474 POTENTIAL.

FT TRANSMEM 486 506 POTENTIAL.  
 FT TRANSMEM 535 555 POTENTIAL.  
 FT TRANSMEM 613 633 POTENTIAL.  
 SQ SEQUENCE 654 AA; 73525 MW; 997BD093F7A907B4 CRC64;  
 Query Match 6.5%; Score 115; DB 1; Length 654;  
 Best Local Similarity 20.6%; Pred. No. 0.61;  
 Matches 78; Conservative 59; Mismatches 93; Indels 148; Gaps 22;  
 QY 23 IMRLYFNNGGKR--IWFSTFETAGPVPVIFIPILFLFSVITRRR-----SNNVGDSTSPRLI- 75  
 Db 52 LFNLEFNLGNKQGIWF-----PLLVSFIVSIGALIITIASYIGVRTSPFLVY 97  
 QY 76 --KPRLL-LIAAIVIGILSGFDNYLYAYGIAYLPVSTAALIIASQLAFIAIFSFVMVKHF 132  
 Db 98 RCKPKIRKLSLIIDILSGIPSVIFG-----LFAQQLSI-----FFRDILKL 140  
 QY 133 TPPT-INAVVLLT--VGAALVGHHTDKPKVHETHKQYITGFLITVAAAV-----MYA 182  
 Db 141 PPLSLNLNVIAMLSFMIPIVISLTNT-----LTYVNDLISVVSLGENKTSAIYK 192  
 QY 183 FI-----LPLVELAYOKA-KQTMSTYTLVLE----- 206  
 Db 193 IIKKEIKPQLTVILTAFARAISETWAVNFVLOSVMYQEVNNRFTSDLTGLGVIST 252  
 QY 207 -----FQILCLLASIVSVIGMFIAGDFKQALPKAREF-----KLGEA 245  
 Db 253 FIFSENGDEQINGVLVIFGIIILVSLNFFAIWSAN-----PKTLERYPFLKKISNF 306  
 QY 246 LFVVAVFSAIWQGFPLGAIGLIP-----STSLVSGIMISVLLPTEVLAVITFYEHEKQ 301  
 Db 307 IYQVW-----W--FIPNNISALFVLDLTSTQSVKKIKVN-----NINE--RSLFFKERLQ 352  
 QY 302 AEKGLSLALSLSLWGFVSVE 319  
 Db 353 S-----VWIKLNYF 362

RESULT 10  
 PAGO\_SALTY  
 ID PAGO\_SALTY STANDARD; PRT; 304 AA.  
 AC O30646;  
 DT 15-JUL-1999 (Rel. 38, Created)  
 DT 15-JUL-1999 (Rel. 38, Last sequence update)  
 DT 28-FEB-2003 (Rel. 41, Last annotation update)  
 DE PAGO protein.  
 GN PAGO OR STM1862.  
 OS Salmonella typhimurium.  
 OC Bacteria; Proteobacteria; Gammaproteobacteria; Enterobacteriales;  
 OC Enterobacteriaceae; Salmonella.  
 OX NCBI\_TaxID=602;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=ATCC 14028S;  
 RX MEDLINE=98380520; PubMed=9712687;  
 RA Gunn J.S., Belden W.J., Miller S.I.;  
 RT "Identification of PhoP-PhoQ activated genes within a duplicated  
 RL region of the Salmonella typhimurium chromosome.";  
 RL Microb. Pathog. 25:77-90(1998).  
 RN [2]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=LT2 / SGSC1412 / ATCC 700720;  
 RX MEDLINE=21534948; PubMed=11677609;  
 RA McClelland M., Sanderson K.E., Spieth J., Clifton S.W., Latreille P.,  
 RA Courtney L., Porwollik S., Ali J., Dante M., Du P., Hou S., Layman D.,  
 RA Leonard S., Nguyen C., Scott K., Holmes A., Grewal N., Mulvaney E.,  
 RA Ryan E., Sun H., Florea L., Miller W., Stoneking T., Nhan M.,  
 RA Waterston R., Wilson R.K.;  
 RT "Complete genome sequence of Salmonella enterica serovar Typhimurium  
 RL LT2.";  
 RL Nature 413:852-856(2001).  
 CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Potential).  
 CC -1- SIMILARITY: Belongs to the eamA transporter family.



```
QY 28 PNNG-----GKRIWFSTFLETAGPPV-----IFI-----PLLFYSYITRRSNNVG 67
Db 31 FGNFGFMDLQGIYLLKFTVDVAGIPAAWAGGIFLVSKLFAAITDPIVGGSSIDYRK--NIG 88
QY 68 DSTSFLLKPRLLIAAVVIGL-----SGFDNVLVAY-----GIAY-----LPV- 106
Db 89 KRGF--RYPYLLIGSVLAVLTVLIFLSPNVSTGKLIYAYAYMTWIGYSFVNIPYG 145
QY 107 STAAIIASOLAFIAFSPFVKHKPTFTINAVLLTVGAALVGMHTETDKPVHETHKQ 166
Db 146 SLGAAMTQNGEDRTSISTFRQIG-----SIGALFITSV--AVMPLLVKFDNPK----- 191
QY 167 YITGFLITV-----AAVYAFIL-----PLVELAYOKAKQTMG-----YTLVLEPQLILCL 213
Db 192 --VGYPVVMGLFAALGVFWFYCYRNCERIIISEAPKEKLTLSVVVTKFTITNKPILLTV 249
QY 214 LASTSVVIGMFIAIGDFOALPKKEAREFKLGEALFYVVAIFSAILWQGFELGAILG----- 268
Db 250 LMTIFS-----SAYNIKSMVLVPAQVNLGNV--ELWAYVNFILIGSFLGVVFLPKLVK 303
QY 269 IFS-----TSSLVSGI-----MISVLLP-----ITEVLAVIFYHEKFOAEKGLSLALS- 315
Db 304 MFGKRTAMIGFISVAADLINFMLPSNVVYFTILASI-----AFIGISIPNGITWAL 356
QY 316 VSYF--YGEIKSGEDK 329
Db 357 VSDIIDYGEWKSGERK 372

RESULT 12
YETK_BACSU
ID_YETK_BACSU STANDARD; PRT; 330 AA.
AC 031540;
AT 15-JUL-1999 (Rel. 38, Created)
DT 15-JUL-1999 (Rel. 38, Last sequence update)
DT 10-OCT-2003 (Rel. 42, Last annotation update)
DE Hypothetical transport protein yetK.
GN YETK OR BSU07210.
OS Bacillus subtilis.
OC Bacteria; Firmicutes; Bacillales; Bacillaceae; Bacillus.
OX NCBI_TaxID=1423;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=168;
RX MEDLINE=98044033; PubMed=9384377;
RA Kunst F., Ogasawara N., Moszer I., Albertini A.M., Alloni G.,
RA Azevedo V., Bertero M.G., Bessieres P., Bolotin A., Borchert S.,
RA Borriello S., Boursier L., Brans A., Braun M., Brignell S.C., Bron S.,
RA Brouillet S., Bruschi C.V., Caldwell B., Capuano V., Carter N.M.,
RA Choi S.K., Codani J.J., Connerston I.F., Cummings N.J., Daniel R.A.,
RA Denizot F., Devine K.M., Dusterhoft A., Ehrlich S.D., Emmerson P.T.,
RA Entian K.D., Errington J., Fabret C., Ferrari E., Foulger D.,
RA Fritz C., Fujita M., Fujita Y., Fuma S., Galizzi A., Galleron N.,
RA Ghim S.Y., Glaeser P., Goffeau A., Gollightly E.J., Grandi G.,
RA Guisepi G., Guy B.J., Haga K., Halech J., Harwood C.R., Henaut A.,
RA Hilbert H., Holsappel S., Hosono S., Hullo M.F., Itaya M., Jones L.,
RA Joris B., Karamata D., Kasahara Y., Klaert-Blanchard M., Klein C.,
RA Kobayashi Y., Koetter P., Koningsstein G., Krogh S., Kumano M.,
RA Kurita K., Lapidus A., Lardinois S., Lauber J., Lazarevic V.,
RA Lee S.M., Levine A., Liu H., Masuda S., Maue C., Medigue C.,
RA Medina N., Mellado R.P., Mizuno M., Moestel D., Nakai S., Noback M.,
RA Noone D., O'Reilly M., Ogawa K., Ogiwara A., Oudega B., Park S.H.,
RA Parro V., Pohl T.M., Portetelle D., Porwollik S., Prescott A.M.,
RA Pressac E., Puig P., Purnelle B., Rapoport G., Rey M., Reynolds S.,
RA Rieger M., Rivolta C., Rocha E., Roche B., Rose M., Sadaie Y.,
RA Sato T., Scanlan E., Schleich S., Schroeter R., Scoffone F.,
RA Sekiguchi J., Sekowska A., Seror S.J., Seror P., Shin B.S., Soldo B.,
RA Sorokin A., Taconi E., Takagi T., Takahashi H., Takemaru K.,
RA Takeuchi M., Takakoshi A., Tanaka T., Terpetra P., Tognoni A.,
RA Tosato V., Uchiyama S., Vandenbol M., Vannier F., Vassarotti A.,
RA Viari A., Wambutt R., Wedler E., Wedler H., Weitzengger T.,
RA Winters P., Wipat A., Yamamoto H., Yamane K., Yasumoto K., Yata K.,
RA Yoshida K., Yoshikawa H.F., Zumstein E., Yoshikawa H., Danchin A.,
```

```
RT "The complete genome sequence of the Gram-positive bacterium Bacillus
RL subtilis."
CC -1- SUBCELLULAR LOCATION: Integral membrane protein (Probable).
CC -1- SIMILARITY: Belongs to the ema transporter family.
CC -----
CC THIS SWISS-PROT entry is copyrighted. It is produced through a collaboration
CC between the Swiss Institute of Bioinformatics and the EMBL outstation -
CC the European Bioinformatics Institute. There are no restrictions on its
CC use by non-profit institutions as long as its content is in no way
CC modified and this statement is not removed. Usage by and for commercial
CC entities requires a license agreement (see http://www.isb-sib.ch/announce/
CC or send an email to license@isb-sib.ch).
CC -----
CC EMBL; Z99107; CAB12540.1; -.
DR PIR; H69798; H69798.
DR Subtilisin; BGI2867; yetK.
DR InterPro; IPR000620; DUF6.
DR Pfam; PF00892; DUF6; 2.
KW Hypothetical protein; Transport; Transmembrane; Complete proteome.
FT TRANSMEM 27 47
FT TRANSMEM 56 76
FT TRANSMEM 90 110
FT TRANSMEM 119 139
FT TRANSMEM 147 167
FT TRANSMEM 176 196
FT TRANSMEM 206 226
FT TRANSMEM 243 263
FT TRANSMEM 270 290
FT TRANSMEM 294 314
FT TRANSMEM 330 AA; 35758 MW; 7C6C65F1EDB276B5 CRC64;
SQ SEQUENCE 330 AA; 6.4%; Score 112.5; DB 1; Length 330;
Query Match 6.4%; Score 112.5; DB 1; Length 330;
Best Local Similarity 19.8%; Pred. No. 0.48;
Matches 68; Conservative 59; Mismatches 127; Indels 89; Gaps 15;
QY 1 MVRALVIINCIILAIAGNCGGLIMRL-YFNGGKRINFST-----FLETAGPPV--- 48
Db 33 LAAAMAIVGSSVVV---GKLMVERIPVFLSSGLRFLIASVLLMLLFCIEKGFFALTKK 88
QY 49 -IFIPILLFSYITRRSNNVGDSTSPFLIKPRLLIAAVIVGILSGFDNLYVAYGIAPVS 107
Db 89 DVFVLLVQSF-----TGVELFSICLL-----YGVQYTTGT 118
QY 108 TAAIIITASQLAFTAFISFFVMVHKFTPTFTINAVLLTVGAALV---GMHTETDKPVHETH 164
Db 119 ESGILSTTPMLIGILSFFLIREKIEKKTLIGILLAVCGVMALNLFAGSQDGT- 173
QY 165 KOYITGFLITVAAAVMYAFILPLVELAYQKAKQTMSTYTLVLEPQLILCLLASIVSVIG-- 222
Db 174 -HALFGNMLIIAAVIGEAFLTLAKL-----LSPHISALAISTFVSLFGFL 218
QY 223 -----MFIAGDPKQALPKEAR-EFKLGEALFYVVAIFSAILWQGFELGIFSTSSL 275
Db 219 FFLFPALFEASSPDYSVPTVLDWSYVLYALFVTVLAF--YLMW---YSG---VTKVPAG 269
QY 276 VSGIMLSVLLPITEVLAVIFYHEKFOAEKGLSLALSLSWG-FVS 317
Db 270 VSGIFTSVLPVSAILVSGVILKEPFEVHFPIGFIACVIGTIGFVT 312

RESULT 13
YB64_SYNY3
ID_YB64_SYNY3 STANDARD; PRT; 349 AA.
AC P73771;
DT 01-NOV-1997 (Rel. 35, Created)
DT 01-NOV-1997 (Rel. 35, Last sequence update)
DT 28-FEB-2003 (Rel. 41, Last annotation update)
DE Hypothetical transport protein sl1164.
GN SL1164.
OS Synechocystis sp. (strain PCC 6803).
OC Bacteria; Cyanobacteria; Chroococcales; Synechocystis.
OX NCBI_TaxID=1148;
```

[1]  
 RN SEQUENCE FROM N.A.  
 RX MEDLINE=97061201; PubMed=8905231;  
 RA Kaneko T., Sato S., Kotani H., Tanaka A., Asamizu E., Nakamura Y.,  
 RA Miyajima N., Hirosewa M., Sugiyama M., Sasanoto S., Kimura T.,  
 RA Hosouchi T., Matsuno A., Muraki A., Nakazaki N., Naruo K., Okumura S.,  
 RA Shimo S., Takeuchi C., Wada T., Watanabe A., Yamada M., Yasuda M.,  
 RA Tabata S.;  
 RT "Sequence analysis of the genome of the unicellular cyanobacterium  
 RT Synchocystis sp. strain PCC6803. II. Sequence determination of the  
 RT entire genome and assignment of potential protein-coding regions.";  
 RL DNA Res. 3:109-136(1996).  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein (Probable).  
 CC -!- SIMILARITY: Belongs to the eamA transporter family.  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; D90909; BAA17823.1; -;  
 DR PIR; S74862; S74862.  
 DR InterPro; IPR000620; DUF6.  
 DR Pfam; PF00892; DUF6; 2.  
 KW Hypothetical protein; Transpos; Transmembrane; Complete proteome.  
 FT TRANSMEM 15 35  
 FT TRANSMEM 53 73  
 FT TRANSMEM 91 111  
 FT TRANSMEM 120 140  
 FT TRANSMEM 147 167  
 FT TRANSMEM 179 199  
 FT TRANSMEM 218 238  
 FT TRANSMEM 248 268  
 FT TRANSMEM 276 296  
 FT TRANSMEM 302 322  
 FT SEQUENCE 349 AA; 38558 MW; DF56413379367376 CRC64;  
 Query Match 6.3%; Score 111.5; DB 1; Length 349;  
 Best Local Similarity 20.2%; Pred. No. 0.6;  
 Matches 70; Conservative 65; Mismatches 117; Indels 95; Gaps 15;  
 QY 88 ILSGDNLY-YAGIAYLPVSTAAIIASQLA--FAIFSPFVHKHPTPTINAVV--- 141  
 DB 4 VFGSGFTLQVHSPYILALAVVIFSAANPVTISRILGLKTYEINGR-NPISFCNVLFVG 62  
 QY 142 -LITVGNALVGMHTETDKPVE---THKOYITGLITVAAAVVAFILPLVELAYQKAK 196  
 DB 63 NLGALGMLITFR--PDWQHKURKUTRQDW---FLTITAILSRALAPGLMPSALEKTN 117  
 QY 197 QT-----MSYTLVLEFQLI-----LCLLASIVSGMFIAGDFKQALPKEA---R 238  
 DB 118 VTNVVLIGRLEPFTLILSILLKISVNWLSMVALISFGVAVTVFVGADPLDMVINF 177  
 QY 239 EFKLGEA-----LFTYVAVF----- 253  
 DB 178 DFLGGSFVAIAAFISAITTILSKQQLQSPVGIPTVYRSLGTFVFFVIAVLIYGFDFP 237  
 QY 254 ----SALIWGGFFL-GAIGLIFSTSSIVSG-----IMISVLPIITEVLAVIFVH----E 298  
 DB 238 MDVSSPLLRWMLIYGAILVVGVQVAVLAKGNASFIQINLASLIVTPILAIIFAYLILLE 297  
 QY 299 KFOAEKGLSALSUNGVSFVYGEIKSGEDKRIQBESEQETQSSIL 345  
 DB 298 TPTNAQYLGILLGLLGLSILFDNLKTAKDKQAKPLNSREAMDNTV 344  
 RESULT 14  
 DUR3 YEAST  
 ID -DUR3 YEAST  
 AC P33413;

DT 01-FEB-1994 (Rel. 28, Created)  
 DT 01-FEB-1995 (Rel. 31, Last sequence update)  
 DT 10-OCT-2003 (Rel. 42, Last annotation update)  
 DE Urea active transporter.  
 GN DUR3 OR YHL016C.  
 OS Saccharomyces cerevisiae (Baker's yeast).  
 OC Eukaryota; Fungi; Ascomycota; Saccharomycotina; Saccharomycetes;  
 OC Saccharomycetales; Saccharomycetaceae; Saccharomycetes.  
 OC NCBI\_TaxID=4932;  
 RN [1]  
 RP SEQUENCE FROM N.A.  
 RC STRAIN=TCY1;  
 RX MEDLINE=93328673; PubMed=8335627;  
 RA Elberry H.M., Majumdar M.L., Cunningham T.S., Sumrada R.A.,  
 RA Cooper T.G.;  
 RA "Regulation of the urea active transporter gene (DUR3) in  
 RT Saccharomyces cerevisiae.";  
 RL J. Bacteriol. 175:4688-4698(1993).  
 CC [2]  
 CC SEQUENCE FROM N.A.  
 RC STRAIN=S288C / AB972;  
 RX MEDLINE=94378003; PubMed=8091229;  
 RA Johnston M., Andrews S., Brinkman R., Cooper J., Ding H., Dover J.,  
 RA Du Z., Favell A., Fulton L., Gattung S., Geisel C., Kirsten J.,  
 RA Kucaba T., Hillier L.W., Jier M., Johnston L., Langston Y.,  
 RA Latreille P., Louis E.J., Macri C., Mardis E., Menezes S., Mouser L.,  
 RA Nham M., Rifkin L., Riles L., St Peter H., Trevaskis E., Vaughan K.,  
 RA Vignati D., Wilcox L., Wohlman P., Waterston R., Wilson R.,  
 RA Vaudin M.;  
 RA "Complete nucleotide sequence of Saccharomyces cerevisiae chromosome  
 RT VIII.";  
 RL Science 265:2077-2082(1994).  
 CC -!- FUNCTION: REQUIRED FOR ACTIVE TRANSPORT OF UREA.  
 CC -!- SUBUNIT: MAY POLYMERIZE.  
 CC -!- SUBCELLULAR LOCATION: Integral membrane protein.  
 CC -!- INDUCTION: BY ALLOPHANATE OR ITS NON-METABOLIZED ANALOG OXALURATE.  
 CC -!- SENSITIVE TO NITROGEN CATABOLITE REPRESSION.  
 CC -!- SIMILARITY: BELONGS TO THE SODIUM:SOLUTE SYMPORTER FAMILY (SSF).  
 CC  
 CC This SWISS-PROT entry is copyright. It is produced through a collaboration  
 CC between the Swiss Institute of Bioinformatics and the EMBL outstation -  
 CC the European Bioinformatics Institute. There are no restrictions on its  
 CC use by non-profit institutions as long as its content is in no way  
 CC modified and this statement is not removed. Usage by and for commercial  
 CC entities requires a license agreement (See http://www.isb-sib.ch/announce/  
 CC or send an email to license@isb-sib.ch).  
 CC  
 CC EMBL; L19875; AAA34582.1; -;  
 DR EMBL; U11582; AAB65069.1; -;  
 DR PIR; S46830; S46830.  
 DR GerMOnline; I39283; -;  
 DR SGD; S0001008; DUR3.  
 DR GO; GO:0005886; C:plasma membrane; IMP.  
 DR GO; GO:0015204; F:urea transporter activity; IMP.  
 DR GO; GO:0015840; F:urea transport; IMP.  
 DR InterPro; IPR001734; Na/solut\_symport.  
 DR Pfam; PF00474; SSF; 1.  
 DR TIGRFAme; TIGR00813; sss; 1.  
 DR PROSITE; PS00456; NA\_SOLUT\_SYMP\_1; FALSE NEG.  
 DR PROSITE; PS00457; NA\_SOLUT\_SYMP\_2; FALSE NEG.  
 DR PROSITE; PS0283; NA\_SOLUT\_SYMP\_3; 1.  
 KW Transport; Transmembrane.  
 FT TRANSMEM 15 35  
 FT TRANSMEM 86 106  
 FT TRANSMEM 131 151  
 FT TRANSMEM 167 187  
 FT TRANSMEM 190 210  
 FT TRANSMEM 234 274  
 FT TRANSMEM 296 316  
 FT TRANSMEM 344 364  
 FT TRANSMEM 404 424  
 FT TRANSMEM 426 446  
 FT TRANSMEM 455 475



us-09-913-767-9.rsp

Mon Sep 27 08:51:17 2004

Search completed: September 23, 2004, 23:50:23  
Job time : 16.9153 secs

GenCore version 5.1.6  
Copyright (c) 1993 - 2004 CompuGen Ltd.

OM protein - protein search, using sw model

Run on: September 23, 2004, 23:33:46 ; Search time 74.4181 Seconds  
(without alignments)  
1509.371 Million cell updates/sec

Title: us-09-913-767-8  
Perfect score: 1798  
Sequence: 1 MNKGLIIINCIIITIGTCGG.....DRQPPELPILPVS DYVA 356

Scoring table: BLOSUM62  
Gapop 10.0 , Gapext 0.5

Searched: 1017041 seqs, 315518202 residues

Total number of hits satisfying chosen parameters: 1017041

Minimum DB seq length: 0  
Maximum DB seq length: 2000000000  
Post-processing: Minimum Match 0%  
Maximum Match 100%  
Listing first 45 summaries

Database : SPREMBL\_25.\*  
1: sp\_archaea.\*  
2: sp\_bacteria.\*  
3: sp\_fungi.\*  
4: sp\_human.\*  
5: sp\_invertebrate.\*  
6: sp\_mammal.\*  
7: sp\_mhc.\*  
8: sp\_organelle.\*  
9: sp\_phase.\*  
10: sp\_plant.\*  
11: sp\_rodent.\*  
12: sp\_virus.\*  
13: sp\_vertebrate.\*  
14: sp\_unclassified.\*  
15: sp\_rvirus.\*  
16: sp\_bacteriap.\*  
17: sp\_archaeap.\*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

## SUMMARIES

Result No.	Score	Query Match	Length	DB ID	Description
1	1798	100.0	356	10 Q9M7R1	Q9m7r1 arabidopsis
2	1792	99.7	356	10 Q9FZ96	Q9fz96 arabidopsis
3	1160	64.5	358	10 P93010	P93010 arabidopsis
4	1160	64.5	358	10 Q94GB1	Q94gb1 arabidopsis
5	1134.5	63.1	351	10 Q9FZ95	Q9fz95 arabidopsis
6	676.5	37.6	399	10 Q8H7P2	Q8h7p2 oryza sativ
7	553	30.8	382	10 Q9SY29	Q9sy29 arabidopsis
8	550	30.6	379	10 Q9LPF6	Q9lpf6 arabidopsis
9	545.5	30.3	344	10 Q49726	Q49726 arabidopsis
10	543	30.2	348	10 Q49725	Q49725 arabidopsis
11	535	29.8	751	10 Q7X7P4	Q7x7p4 oryza sativ
12	533.5	29.7	1128	10 Q49724	Q49724 arabidopsis
13	532	29.6	372	10 Q8S170	Q8s170 oryza sativ
14	531.5	29.6	315	10 Q9ZUH3	Q9zuh3 arabidopsis
15	526.5	29.3	377	10 Q8RY74	Q8ry74 arabidopsis
16	494.5	27.5	358	10 Q49722	Q49722 arabidopsis

17	473.5	26.3	361	10 Q8RY83	Q8ry83 arabidopsis
18	473.5	26.3	432	10 Q9LDT1	Q9ldt1 arabidopsis
19	465.5	25.9	384	10 Q7XN26	Q7xn26 oryza sativ
20	445	24.7	358	10 Q9FLA4	Q9fla4 arabidopsis
21	327	18.2	393	10 Q9FXH5	Q9fxh5 arabidopsis
22	313	17.4	389	10 Q9SX90	Q9sx90 arabidopsis
23	301.5	16.8	232	10 Q8H7F0	Q8h7f0 arabidopsis
24	287	16.0	390	10 Q9C508	Q9c508 arabidopsis
25	282.5	15.7	392	10 Q9SX93	Q9sx93 arabidopsis
26	282.5	15.7	394	10 Q9C654	Q9c654 arabidopsis
27	271	15.1	398	10 Q9C655	Q9c655 arabidopsis
28	258	14.3	383	10 Q04508	Q04508 arabidopsis
29	242	13.5	387	10 Q9LQZ0	Q9lqz0 arabidopsis
30	219.5	12.2	149	10 Q94JW7	Q94jw7 arabidopsis
31	156.5	8.7	322	5 Q81BT1	Q81bt1 plasmodium
32	142.5	7.9	280	17 Q977Z4	Q977z4 thermoplaem
33	140	7.8	425	3 Q9P3N3	Q9p3n3 neurospora
34	138	7.7	368	10 Q9SUF0	Q9suf0 arabidopsis
35	138	7.7	519	3 Q9P6J7	Q9p6j7 schizosacch
36	136.5	7.6	285	17 Q9HI53	Q9hi53 thermoplaem
37	135	7.5	644	10 Q7X7A0	Q7x7a0 oryza sativ
38	134	7.5	376	10 Q8LR17	Q8lr17 oryza sativ
39	132.5	7.4	299	2 Q8KH10	Q8kh10 bacillus th
40	132	7.3	311	16 Q9KFT7	Q9kft7 bacillus ha
41	131.5	7.3	300	16 Q8XRZ6	Q8xrz6 ralstonia s
42	130.5	7.3	303	16 Q8ISM6	Q8ism6 bacillus an
43	130.5	7.3	414	10 Q8H184	Q8h184 arabidopsis
44	130.5	7.3	414	10 Q84WH7	Q84wh7 arabidopsis
45	129.5	7.2	322	16 Q8RE82	Q8re82 fusbacteri

## ALIGNMENTS

## RESULT 1

Q9M7R1 PRELIMINARY; PRT; 356 AA.  
ID Q9M7R1  
AC Q9M7R1;  
DT 01-OCT-2000 (TREMBLrel. 15, Created)  
DT 01-OCT-2000 (TREMBLrel. 15, Last sequence update)  
DE 01-JUN-2003 (TREMBLrel. 24, Last annotation update)  
DE Purine permease.  
DS PUP1.  
OS Arabidopsis thaliana (Mouse-ear cress).  
GN Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosidb II; Brassicales; Brassicaceae; Arabidopsi.  
OX NCBI\_TaxID=3702;  
RN [1]  
RP SEQUENCE FROM N.A.  
RC STRAIN=cv. Columbia;  
RX MEDLINE=20129770; PubMed=10662864;  
RA Gillissen B., Burkle L., Andre B., Kuhn C., Rentsch D., Brandl B.,  
RA Frommer W.B.;  
RT "A new family of high-affinity transporters for adenine, cytosine, and  
RT purine derivatives in Arabidopsis.";  
RL Plant Cell 12:291-300(2000).  
DR EMBL; AF078531; AAF64547.1;  
DR GO; GO:0016020; C:membrane; IEA.  
DR InterPro; IPR004853; DUF250.  
DR InterPro; IPR000620; DUF6.  
DR Pfam; PF03151; DUF250; 1.  
DR Pfam; PF00892; DUF6; 1.  
SQ SEQUENCE 356 AA; 39250 MW; 0A38267726081261 CRC64;

Query Match 100.0%; Score 1798; DB 10; Length 356;  
Best Local Similarity 100.0%; Pred. No. 5,Se-133;  
Matches 356; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

QY 1 MNKGLIIINCIIITIGTCGGPLLTRLYFTNGGKRIWFMFSLTAGPPIILIPLLVFLSR 60  
DB 1 MNKGLIIINCIIITIGTCGGPLLTRLYFTNGGKRIWFMFSLTAGPPIILIPLLVFLSR 60

```

QY 61 RRGNNPNAENKRTKLFMETPLFIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIG 120
DB 61 RRGNNPNAENKRTKLFMETPLFIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIG 120
QY 121 TOLAFNALFAFLLVKQKFTPFSSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
DB 121 TOLAFNALFAFLLVKQKFTPFSSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
QY 181 VVAALLYAFILPLVELTYKKARQETFPFLVLEIQMWMLAAATFFCVIGMFTVGDVKVIAR 240
DB 181 VVAALLYAFILPLVELTYKKARQETFPFLVLEIQMWMLAAATFFCVIGMFTVGDVKVIAR 240
QY 241 EAREFKIGSVFYFVYALIVITGIWQGFELGAIIVFCASSLASGVLISVLLPVTVEFAVV 300
DB 241 EAREFKIGSVFYFVYALIVITGIWQGFELGAIIVFCASSLASGVLISVLLPVTVEFAVV 300
QY 301 CFREKFOAEKGVSLLSLWGFVSFYGFYFGEFSGKGVVDKPPQPPETELPLPVSDDYA 356
DB 301 CFREKFOAEKGVSLLSLWGFVSFYGFYFGEFSGKGVVDKPPQPPETELPLPVSDDYA 356

```

```

RESULT 2
Q9FZ96 PRELIMINARY; PRT; 356 AA.
ID Q9FZ96;
AC Q9FZ96;
DT 01-MAR-2001 (Tremblrel. 16, Created)
DT 01-MAR-2001 (Tremblrel. 16, Last sequence update)
DT 01-JUN-2003 (Tremblrel. 24, Last annotation update)
DE Purine permease.
GN F3H9.22 OR ATIG28230/F3H9 10 OR ATIG28230.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F., S.,
RA Altati H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chin C., Chioi J., Choi E., Gonzalez A.,
RA Howng B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharey N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vaysberg M., Yu G., Ecker J., Theologis A., Davis R.W.;
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Nakajima M., Enju A., Kamiya A., Narusaka M., Carninci P., Kawai J.,
RA Hayashizaki Y., Shinozaki K.;
RA "Arabidopsis thaliana full-length cDNA.";
RT Submitted (NOV-2002) to the EMBL/GenBank/DBJ databases.
[3]
RN SEQUENCE FROM N.A.
RA Yamada K., Chan M.M., Chang C.H., Dale J.M., Hsuan V.W., Lee J.M.,
RA Onodera C.S., Quach H.L., Tang C., Toriumi M., Wong C., Wu H.C.,
RA Yu G., Yuan S., Carninci P., Chen H., Cheuk R., Hayashizaki Y.,
RA Ishida J., Jones T., Kamiya A., Kawai J., Kim C.J., Narusaka M.,
RA Nguyen M., Palm C.J., Sakurai T., Satou M., Seki M., Shinn P.,
RA Southwick A., Tripp M.G., Wu T., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RT "Arabidopsis Open Reading Frame (ORF) Clones.";
RL Submitted (MAR-2003) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC021044; AAF98432.1; -.
DR EMBL; AK117664; BAC42317.1; -.
DR EMBL; BT005504; AAC63924.1; -.
DR PIR; D86408; D86408.
DR GO; GO:0016020; C:membrane; IEA.
DR InterPro; IPR004853; DUF250.
DR Pfam; PF03151; DUF250; DUF6.
DR Pfam; PF00892; DUF6; 1.
SQ SEQUENCE 356 AA; 39280 MW; FA2701E954CF797B CRC64;

```

```

Query Match 99.7%; Score 1792; DB 10; Length 356;
Best Local Similarity 99.7%; Pred. No. 1.6e-132;
Matches 355; Conservative 0; Mismatches 1; Indels 0; Gaps 0;

QY 1 MKNGLIINIICIIITIGTCGGPLLTRLYFTNGGKRIWFMSFLTAGFPPIILIPLLVSFLSR 60
DB 1 MKNGLIINIICIIITIGTCGGPLLTRLYFTNGGKRIWFMSFLTAGFPPIILIPLLVSFLSR 60
QY 61 RRGNNPNAENKRTKLFMETPLFIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIG 120
DB 61 RRGNNPNAENKRTKLFMETPLFIASIVIGLTLGDLNLYSYGLAYLPVSTSSLIIG 120
QY 121 TOLAFNALFAFLLVKQKFTPFSSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
DB 121 TOLAFNALFAFLLVKQKFTPFSSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMT 180
QY 181 VVAALLYAFILPLVELTYKKARQETFPFLVLEIQMWMLAAATFFCVIGMFTVGDVKVIAR 240
DB 181 VVAALLYAFILPLVELTYKKARQETFPFLVLEIQMWMLAAATFFCVIGMFTVGDVKVIAR 240
QY 241 EAREFKIGSVFYFVYALIVITGIWQGFELGAIIVFCASSLASGVLISVLLPVTVEFAVV 300
DB 241 EAREFKIGSVFYFVYALIVITGIWQGFELGAIIVFCASSLASGVLISVLLPVTVEFAVV 300
QY 301 CFREKFOAEKGVSLLSLWGFVSFYGFYFGEFSGKGVVDKPPQPPETELPLPVSDDYA 356
DB 301 CFREKFOAEKGVSLLSLWGFVSFYGFYFGEFSGKGVVDKPPQPPETELPLPVSDDYA 356

RESULT 3
P93010 PRELIMINARY; PRT; 356 AA.
ID P93010;
AC P93010;
DT 01-MAY-1997 (Tremblrel. 03, Created)
DT 01-MAY-1997 (Tremblrel. 03, Last sequence update)
DT 01-OCT-2003 (Tremblrel. 25, Last annotation update)
DE AT2G33750 protein.
GN AT2G33750.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eucosids II; Brassicales; Brassicaceae; Arabidopsis.
OX NCBI_TaxID=3702;
[1]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Renning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Creasy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Freus D., Nierman W.C., White O., Bisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RT "Sequence and analysis of chromosome 2 of the plant Arabidopsis
thaliana.";
RL Nature 402:761-768(1999).
[2]
RN SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; U78721; AAC69140.1; -.
DR PIR; B84749; B84749.
DR InterPro; IPR004853; DUF250.
DR Pfam; PF03151; DUF250; 1.
SQ SEQUENCE 356 AA; 39201 MW; CD5BCAE31BE63E6 CRC64;

Query Match 64.5%; Score 1160; DB 10; Length 356;
Best Local Similarity 62.9%; Pred. No. 5.3e-83;
Matches 231; Conservative 45; Mismatches 69; Indels 22; Gaps 4;

```



```

QY 1 MNGLIIINCIITIGTCGGPLLTRLYFTNGGRIWFMFSLTAGPPIILPLLVSLR 60
DB 1 MKTVLIINCIIFLAIGNCGPLMRLYFQNGGERIWFPSPLOTVGCPDLPFLLSFLR 60
QY 61 RRGNNPNAENKRTKLFMEYPLFTASIVIGLITGLDNLVSYGLAYLPVSTSLIIG 120
DB 61 RR-----CLEQETTPFFLMKPLFLFAIIVGLLVGDFDNLVSYGLAYIPVSTSLIIS 114
QY 121 TQAFNALFAFLVVKQKFTPFPSINAVLLTVGIGLALHSDGDKPKAKESKKEVYVGFMT 180
DB 115 AQLGFTALFAFFVMKQKFTPFPSINAVLLTVGIGLALHSDGDKPKAKESKKEVYVGFMT 174
QY 181 VVAALLYAFILPLVELTYKKARQETIPPLVLEIQMVMCLAAATFFCVIGMFIIVGDFK 236
DB 175 LGAALLYGFLPLVELSYKSGORITVTLALEFQWMLCFAATCVCLVGMALAGDFKVKHA 234
QY 237 -----VIAREAREFKIGSVFYALIVITIGIWOQFFLGAIGVFCASSLASGVLSV 289
DB 235 LFTFKNRVIAAGEARDFKLGESL-YVVIVFTAIWQFAFFVGAIGLIFCASSLVSGIMVSA 293
QY 290 LLPVTEVFAVCPREKFOAEKGVSLLSLWGFVSFYFGEFKSKKVVDKQPPELPIIL 349
DB 294 LLPVTVLAVICFOERQAGKVALALSLWGSVSFYFGQVKS-----EETKAQDTQLSQL 349
QY 350 PVS DYVA 356
DB 350 PVT DYVA 356

RESULT 4
Q94GB1 PRELIMINARY; PRT; 358 AA.
ID Q94GB1 AC Q94GB1;
DT 01-DEC-2001 (T-EMBLrel. 19, Created)
DT 01-DEC-2001 (T-EMBLrel. 19, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Putative purine permease (fragment).
GN PUP2.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA STRAIN=cv. C 24;
RC MEDLINE=20129770; PubMed=10662864;
RX Gillissen B., Burkle L., Andre B., Kuhn C., Rentsch D., Brandl B.,
RA Frommer W.B.;
RT "A new family of high-affinity transporters for adenine, cytosine, and
RT purine derivatives in Arabidopsis.";
RL Plant Cell 12:291-300(2000).
DR EMBL: AF078532; AAK61813.1;
DR InterPro: IPR004853; DUF250.
DR Pfam: PF03151; DUF250; 1.
FT NON TER 358 358
SQ SEQUENCE 358 AA; 39461 MW; 69709A0EA6ECD6A8 CRC64;

Query Match 64.5%; Score 1160; DB 10; Length 358;
Best Local Similarity 62.9%; Pred. No. 5.4e-83;
Matches 231; Conservative 45; Mismatches 69; Indels 22; Gaps 4;

QY 1 MNGLIIINCIITIGTCGGPLLTRLYFTNGGRIWFMFSLTAGPPIILPLLVSLR 60
DB 3 MKTVLIINCIIFLAIGNCGPLMRLYFQNGGERIWFPSPLOTVGCPDLPFLLSFLR 62
QY 61 RRGNNPNAENKRTKLFMEYPLFTASIVIGLITGLDNLVSYGLAYLPVSTSLIIG 120
DB 63 RR-----CLEQETTPFFLMKPLFLFAIIVGLLVGDFDNLVSYGLAYIPVSTSLIIS 116
QY 121 TQAFNALFAFLVVKQKFTPFPSINAVLLTVGIGLALHSDGDKPKAKESKKEVYVGFMT 180
DB 117 AQLGFTALFAFFVMKQKFTPFPSINAVLLTVGIGLALHSDGDKPKAKESKKEVYVGFMT 176

```

```

QY 181 VVAALLYAFILPLVELTYKKARQETIPPLVLEIQMVMCLAAATFFCVIGMFIIVGDFK 236
DB 177 LGAALLYGFLPLVELSYKSGORITVTLALEFQWMLCFAATCVCLVGMALAGDFKVKHA 236
QY 237 -----VIAREAREFKIGSVFYALIVITIGIWOQFFLGAIGVFCASSLASGVLSV 289
DB 237 LFTFKNRVIAAGEARDFKLGESL-YVVIVFTAIWQFAFFVGAIGLIFCASSLVSGIMVSA 295
QY 290 LLPVTEVFAVCPREKFOAEKGVSLLSLWGFVSFYFGEFKSKKVVDKQPPELPIIL 349
DB 296 LLPVTVLAVICFOERQAGKVALALSLWGSVSFYFGQVKS-----EETKAQDTQLSQL 351
QY 350 PVS DYVA 356
DB 352 PVT DYVA 358

RESULT 5
Q9FZ95 PRELIMINARY; PRT; 351 AA.
ID Q9FZ95 AC Q9FZ95;
DT 01-MAR-2001 (T-EMBLrel. 16, Created)
DT 01-MAR-2001 (T-EMBLrel. 16, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE F3H9.12 protein.
GN F3H9.12.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosid II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Federspiel N.A., Palm C.J., Conway A.B., Conn L., Hansen N.F.,
RA Altafi H., Nguyen M., Lam B., Southwick A., Miranda M., Brooks S.,
RA Buehler E., Chao Q., Chiu J., Choi E., Gonzalez A.,
RA Howng B., Johnson-Hopson C., Khan S., Kim C., Koo T., Lee J.M.,
RA Lenz C., Liu A., Liu S., Mukharasy N., Pham P., Sakano H., Shinn P.,
RA Toriumi M., Vaynsberg M., Yu G., Ecker J., Theologis A., Davis R.W.,
RA Submitted (AUG-2000) to the EMBL/GenBank/DBJ databases.
RL EMBL: AC021044; AAF98433.1;
DR PIR: C86408; C86408.
DR InterPro: IPR004853; DUF250.
DR Pfam: PF03151; DUF250; 1.
SQ SEQUENCE 351 AA; 38871 MW; 992F1704CB76ED70 CRC64;

Query Match 63.1%; Score 1134.5; DB 10; Length 351;
Best Local Similarity 59.9%; Pred. No. 5.2e-81;
Matches 214; Conservative 66; Mismatches 66; Indels 11; Gaps 4;

QY 1 MNGLIIINCIITIGTCGGPLLTRLYFTNGGRIWFMFSLTAGPPIILPLLVSLR 60
DB 1 MKTVLIINCIIFLAIGNCGPLMRLYFQNGGERIWFPSPLOTVGCPDLPFLLSFYTR 60
QY 61 RRGNNPNAENKRTKLFMEYPLFTASIVIGLITGLDNLVSYGLAYLPVSTSLIIG 120
DB 61 RRSNNVGD-----TSFLIKPRLIAAVIVGLSGFDNLVYAGIYLPVSTALIIA 114
QY 121 TQAFNALFAFLVVKQKFTPFPSINAVLLTVGIGLALHSDGDKPKAKESKKEVYVGFMT 180
DB 115 SQLAFIAIFSFMVHKHFTPTINAVLLTVGAAVLGMHTETDKPVHETHKQVITGLIT 174
QY 181 VVAALLYAFILPLVELTYKKARQETIPPLVLEIQMVMCLAAATFFCVIGMFIIVGDFKVIAR 240
DB 175 VAAVMYAFILPLVELAYQAKQTMSTYLLVLEQLICLLASIVSGIMFIIVGDFKALPK 234
QY 241 EAREFKIGSVFYALIVITIGIWOQFFLGAIGVFCASSLASGVLSVLPVTEFAV 300
DB 235 EAREFKIGALP-YVAVFSAIIWQFFLGAIGLIFSTSLVSGIMFIIVGDFKALPK 293
QY 301 CFREKFOAEKGVSLLSLWGFVSFYFGEFKSG--KKVVDKQPPELPIIL--PVSD 353

```



```
DR GO: GO:0016020; C.membrane; IEA.
DR InterPro; IPR004853; DUF250.
DR RA Deng J.M., Hayashizaki Y., Heuan V.W., Lee J.M., Ishida J., Kamiya A.,
DR Kawai J., Kim C.J., Naruaka M., Onodera C.S., Quach H.L., Sakurai T.,
DR Satou M., Seki M., Shinn P., Tang C.C., Torumi M., Wong C., Wu H.C.,
DR Yamada K., Yu G., Yuan S., Shinzaki K., Ecker J., Theologis A.,
DR Davis R.W.;
RA Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC020576; AAF78258.1; -.
DR EMBL; AY062782; AAL32860.1; -.
DR EMBL; BT003358; ABO29976.1; -.
DR PIR; D96506; D96506.
DR InterPro; IPR004853; DUF250.
DR Pfam; PF03151; DUF250; 1.
KW Hypothetical protein.
SQ SEQUENCE 382 AA; 42320 MW; E030A23A8C934ACA CRC64;

Query Match 30.8%; Score 553; DB 10; Length 382;
Best Local Similarity 36.3%; Pred. No. 2.1e-35;
Matches 120; Conservative 65; Mismatches 132; Indels 14; Gaps 4;

QY 5 LIINCIIITIGTCGGPGLLTRLYFTNGGRIWFMSPFLSTAGTPIILIPLVSLRRGN 64
DB LLIIVTFLLFFGSIASSLAKYIFVGGSSRWSTWQSSAGPFLLLILYIPHYVLKTTT 86
QY 65 RNPNAENKRTKFLMETPLFIASIVIGLLTGLDNLVSYGLAYLPVSTSSLIIGTOLA 124
DB RRP-----PFTFTRH---LFSVLIGLVLFNPLFSWGSYLPVSTSSLLTQLV 136
QY 125 FNALPFLVVKQFTPFPSINAVLLTVGIGILAHSDGDKPAKESKEYVVGFLMTVAA 184
DB FTLLSRIIVKQITPSNLNLCVLLTSSVLLALDSSDKPKSGLTTKYFIVGVSTIGAG 196
QY 185 LLYAFILPLVELTYKKARQETPFLVLEIQVMWCLAATPCFVGMFVGDPKVIAREARE 244
DB LFLALPLVTEKLY---RTVYCYAMVEVOLVMEFAATVFPATIGMACGGGFKEMVKEANH 253
QY 245 FKIGGSVFFYALIVITGII-WQGFLLGAIGIVFCASSLASGVLISVLLPTEFVAVVCPR 303
DB VFTKGTFTTWPAILANVTWQLSPATSGMVLTSIGITGICHTALLANVIGGVAYG 313
QY 304 EKPAEKGVSLLSLMGFSVYFYGFEPKSGKK 334
DB 314 DVEFGVKIVSTLVCINGFSGSYTYGMYMKMK 344

RESULT 8
Q9LPP6 PRELIMINARY; PRT; 379 AA.
AC Q9LPP6;
DT 01-OCT-2000 (TrEMBLrel. 15, Created)
DT 01-OCT-2000 (TrEMBLrel. 15, Last sequence update)
DT 01-JUN-2003 (TrEMBLrel. 24, Last annotation update)
DE T12C22.2 protein (Hypothetical protein)
GN T12C22.2 OR ATIG44750, T12C22.2 OR ATIG44750.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Yu G., Vaysberg M., Sakano H., Lee J., Lenz C., Pham C., Toriumi M.,
RA Liu S., Chiu J., Choi E., Chung M., Gonzalez A., Howing B.,
RA Liu A., Altaji H., Brooks S., Buehler E., Chao Q., Conn L., Conway A.,
RA Hansen N., Johnson-Hopson C., Khan S., Kim C., Lam B., Miranda M.,
RA Nguyen M., Palm C., Shinn P., Southwick A., Davis R., Ecker J.,
RA Federspiel N., Theologis A.;
RT "The sequence of BAC T12C22 from Arabidopsis thaliana chromosome 1.";
RL Submitted (JUN-2000) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Southwick A., Karlin-Neumann G., Nguyen M., Lam B., Miranda M.,
RA Palm C.J., Bowser L., Jones T., Banh J., Carninci P., Chen H.,
RA Cheuk R., Chung M.K., Hayashizaki Y., Ishida J., Kamiya A., Kawai J.,
RA Kim C., Lin J., Liu S.X., Narusaka M., Pham P.K., Sakano H.,
RA Sakurai T., Satou M., Seki M., Shinn P., Yamada K., Shinzaki K.,
RA Ecker J., Theologis A., Davis R.W.;
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.

RA Southwick A., Nguyen M., Tripp M., Palm C.J., Jones T., Wu T.,
RA Carninci P., Chen H., Cheuk R., Chan M.M., Chang C.H., Dale J.M.,
RA Deng J.M., Hayashizaki Y., Heuan V.W., Lee J.M., Ishida J., Kamiya A.,
RA Kawai J., Kim C.J., Naruaka M., Onodera C.S., Quach H.L., Sakurai T.,
RA Satou M., Seki M., Shinn P., Tang C.C., Torumi M., Wong C., Wu H.C.,
RA Yamada K., Yu G., Yuan S., Shinzaki K., Ecker J., Theologis A.,
RA Davis R.W.;
RL Submitted (JAN-2003) to the EMBL/GenBank/DBJ databases.
RL EMBL; AC020576; AAF78258.1; -.
DR EMBL; AY062782; AAL32860.1; -.
DR EMBL; BT003358; ABO29976.1; -.
DR PIR; D96506; D96506.
DR InterPro; IPR004853; DUF250.
DR Pfam; PF03151; DUF250; 1.
KW Hypothetical protein.
SQ SEQUENCE 379 AA; 41289 MW; F1ED85A70FAD7001 CRC64;

Query Match 30.8%; Score 550; DB 10; Length 379;
Best Local Similarity 34.4%; Pred. No. 3.5e-35;
Matches 119; Conservative 74; Mismatches 133; Indels 20; Gaps 4;

QY 5 LIINCIIITIGTCGGPGLLTRLYFTNGGRIWFMSPFLSTAGTPIILIPLVSLRRGN 64
DB LLIIVTFLLFFGSIASSLAKYIFVGGSSRWSTWQSSAGPFLLLILYIPHYVLKTTT 99
QY 65 RNPNAENKRTKFLMETPLFIASIVIGLLTGLDNLVSYGLAYLPVSTSSLIIGTOLA 124
DB RRP-----PFTFTRH---LFSVLIGLVLFNPLFSWGSYLPVSTSSLLTQLV 153
QY 125 FNALPFLVVKQFTPFPSINAVLLTVGIGILAHSDGDKPAKESKEYVVGFLMTVAA 184
DB FTLLSRIIVKQITPSNLNLCVLLTSSVLLALDSSDKPKSGLTTKYFIVGVSTIGAG 213
QY 185 LLYAFILPLVELTYKKARQETPFLVLEIQVMWCLAATPCFVGMFVGDPKVIAREARE 244
DB LFLALPLVTEKLY---RTVYCYAMVEVOLVMEFAATVFPATIGMACGGGFKEMVKEANH 273
QY 245 FKIGGSVFFYALIVITGII-WQGFLLGAIGIVFCASSLASGVLISVLLPTEFVAVVCFRE 304
DB VFTKGTFTTWPAILANVTWQLSPATSGMVLTSIGITGICHTALLANVIGGVAYG 332
QY 305 EKPAEKGVSLLSLMGFSVYFYGFEPKSGKKV-----VDKPPPP 342
DB 333 KMSGVKIMAMLIAMGFSVYVYQNHIDDLKVRQAOQAQAGRVEPP 378

RESULT 9
O49726 PRELIMINARY; PRT; 344 AA.
ID O49726;
AC O49726;
DT 01-JUN-1998 (TrEMBLrel. 06, Created)
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T9A21.70 OR AT4G18220.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsi.
OX NCBI_TaxID=3702;
[1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Droat L., Hall C., Hudson S., Ridley P.,
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
[2]
RP SEQUENCE FROM N.A.
RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
[3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
```

Best Local Similarity 34.4%, Pred. No. 1.1e-34;  
Matches 116; Conservative 69; Mismatches 130; Indels 22; Gaps 47

QY 17 TCGGPELLTRLYFTNGGKRIWFMSFLSTAGFPILILPLLVSFLRRRRGNRNNNAENKRT 76  
Db 19 TNGG---SGVYDNGSKWLATVQVGPVLLPYVILSKFTHATTDROGKRTSPNRV 75  
QY 77 KLFMETPFTIASIVIGLLTGLDNYLYSGLAYLPVSTSLIIGTOLAFAALFAFLVVKO 136  
Db 76 LVY-----VVLGLVAGDCYLSIGLLYLPVSTYSLICASQLAFNAFFSYFLNSQ 125  
QY 137 KTFPSINAVLLTVGIGILALHSDGDKPAKESKKEYVGVFPLMTVVAALLYAFILPILVEL 196  
Db 126 KUTPILNSLELLTISLLAFNNEETDSTKVTGKGVKGFICTVAASAGYGLVLSLQOL 185  
QY 197 TYKARQBITPPLVLEIQMVCLAAATFFCVIGMIVGDFKVIAREAREFKIGGSVFFYAL 256  
Db 186 AFLKVLKQNFSEVMDMIYVSLVASCVSVVGLFASSEWKTLSSEMDSNYK-HGKVSIMN 244  
QY 257 IVITGIIWQGFPLGAIGIVFCASSLASGVLISVLLPVTVEPAVVCFFKFOAEGVSLLL 316  
Db 245 LVWTAVTAVTQVFSIGTGLIFELSLFSNAISVIGLVPVPIILAVIIFHDKNGLKVISMIL 304  
QY 317 SLWGFVSFYFGFKSGKKV-----VDKQPDPETE 345  
Db 305 AIWGFYSYVYQYLDKKNLKNHEITTTTSPDPEAE 341

RESULT 11  
Q7X7P4 PRELIMINARY; PRT; 751 AA.

AC 07X7P4;  
DT 01-OCT-2003 (TrEMBLrel. 25, Created)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE OSJNB0013K16.25 protein (OSJNB0016D16.5 protein).  
GN OSJNB0013K16.25 OR OSJNB0016D16.5.  
OS Oryza sativa (Rice).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;  
OC Ehrhartoideae; Oryzeae; Oryza.  
OC NCBI\_TaxID=4530;  
RX [1]  
RN SEQUENCE FROM N.A.  
RP Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,  
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,  
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,  
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,  
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,  
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;  
RL Submitted (DEC-2001) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE FROM N.A.  
RP Han B., Feng Q., Huang Y.C., Li Y., Zhu J.J., Zhao Q., Hu X.,  
RA Liu Y.L., Mu J., Yu Z., Chen L., Fan D.L., Weng Q.J., Zhang L.,  
RA Lu Y.Q., Yu S.L., Liu X.H., Lu T.T., Zhang Y.J., Lu Y., Li C., Li T.,  
RA Zhang Y., Hu H., Jia P.X., Qian Y.M., Ying K., Zhou B., Chen Z.H.,  
RA Hao P., Zhang L., Wu M., Zhang R.Q., Guan J.P., Fu G., Wang S.Y.,  
RA Ren S.X., Lv G., Lin W., Gu W.Q., Zhu G.F., Tu Y.F., Jia J., Yin H.F.,  
RA Zhang Y., Cai Z., Chen J., Kang H., Chen X.Y., Shao C.Y., Sun Y.,  
RA Hu Q.P., Zhang X.L., Zhang W., Wang L.J., Ding C.W., Sheng H.H.,  
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
RA Gu J.L., Chen S.T., Ni L., Zhu F.H., Hong G.F.;  
RL Submitted (SEP-2001) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL662957; CAB03776.1; -.  
DR EMBL; AL60658; CAB04314.1; -.  
SQ SEQUENCE 751 AA; 79987 MW; 714C08319D7B5FD1 CRC64;

Query Match 30.3%; Score 545.5; DB 10; Length 344;  
Best Local Similarity 35.1%, Pred. No. 7.2e-35;  
Matches 120; Conservative 69; Mismatches 142; Indels 11; Gaps 2;

QY 12 ILTIGTCGPELLTRLYFTNGGKRIWFMSFLSTAGFPILILPLLVSFLRRRRGNRNNNAE 71  
Db 7 VQVIGQSVATILGRLYYENGSKWLATVQVGPVLLPYVILSVKTHITTDROGKRTS 66  
QY 72 NKRKTKFLMETPPLFIASIVIGLLTGLDNYLYSGLAYLPVSTSLIIGTOLAFAALFAF 131  
Db 67 LNRNALV-----IVLGLVGAACDYLSIGLLYLPVSTYSLICASQLAFNAFFSY 116  
QY 132 LLVKOKFTFPSINAVLLTVGIGILALHSDGDKPAKESKKEYVGVFPLMTVVAALLYAFIL 191  
Db 117 LNSQKLPILILNSILFLITISLLAFNNEESDKVTGKGVKGFICTVAGASAGFGLL 176  
QY 192 PIVELYTKARQBITPPLVLEIQMVCLAAATFFCVIGMIVGDFKVIAREAREFKIGGSV 251  
Db 177 SLQQLAFKRVLKOTFSEVINMIYMSLVASCVSVVGLFASSEWKTLSSEMDSNYK-LKV 235  
QY 252 FYVALIVITGIIWQGFPLGAIGIVFCASSLASGVLISVLLPVTVEPAVVCFFKFOAEGK 311  
Db 236 SYVMNLVWTAVTAVTQVFSIGTGLIFELSLFSNAISVIGLVPVPIILAVIIFHDKNGLKV 295  
QY 312 VSLLSLWGFVSFYFGFKSGKKVVDKQPDPETELPILPVS 353  
Db 296 ISMILAINGFVSYVYQYLDKKNLKNHEITTTTSPDPEAE 337

RESULT 10  
O49725 PRELIMINARY; PRT; 348 AA.

AC O49725;  
DT 01-JUN-1998 (TrEMBLrel. 06, Created)  
DT 01-JUN-1998 (TrEMBLrel. 06, Last sequence update)  
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)  
DE Hypothetical protein.  
GN T9A21.60 OR AT4G18210.  
OS Arabidopsis thaliana (Mouse-ear cress).  
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;  
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;  
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.  
OC NCBI\_TaxID=3702;  
[1]  
RN SEQUENCE FROM N.A.  
RP Bevan M., Murphy G., Drost L., Hall C., Hudson S., Ridley P.,  
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;  
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.  
[2]  
RN SEQUENCE FROM N.A.  
RP Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;  
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;  
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
[3]  
RN SEQUENCE FROM N.A.  
RP EU Arabidopsis sequencing project;  
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.  
DR EMBL; AL021713; CAB16793.1; -.  
DR EMBL; AL61548; CAB78823.1; -.  
DR PIR; T04923; T04923.  
DR InterPro; IPR004853; DUF250.  
DR Pfam; PF03151; DUF250; 1.  
KW Hypothetical protein.  
SQ SEQUENCE 348 AA; 38556 MW; 354CB8D6D5547786 CRC64;

Query Match 30.2%; Score 543; DB 10; Length 348;  
Best Local Similarity 33.9%, Pred. No. 1e-33;  
Matches 112; Conservative 74; Mismatches 130; Indels 14; Gaps 3

```

QY 5 LIINCIIITGCGPLTRLYFTNGGKRIFWMSFLSTAGFPILIIPLIVSFLSRRGN 64
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 49 MVAVDVFFLIAGQTSATLLGRYYTQGRSKWISAFVRTAGFPILFFTLFFP -----PS 102
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
QY 65 RPNNAENKRTKFLMEPLFTASIVIGLTLGDNLYSYGLAYLPVSTSSLIIGTOLA 124
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  103 KSPSSCNTFPMK-----LAVIYVGLITIAADDMYTGGLKYLIPASTYSILCASQLA 155
QY 125 FNAFLAVLKQKFTPFSSINAVLLTVGIGILALHSDGDKPAKESKKEYVVGFLMTVAA 184
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  156 ENVVPSVLSNQKVTPLIFNSVLLTMSASLIGSVKESQGVGVSGKYLGLFVLTIGAS 215
QY 185 LLYAFILPLVELTYKKARQBITPLVLEIQMVNCLAAATFCVGMIFVDFKVIAREARE 244
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  216 CTYSLIALMLQTLFTETIIKKHTFSAVLNMQIYALVATRAASVVLGFASGEWRSLRGEWMA 275
QY 245 EKIGGSVFYVALIVITGIITWQGFGLGAGIVFCASSLASGVLSVLLPVTVEFVAVVCFRE 304
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  276 FR-SGQSYMILLMAVSNQVANIGVLGIFESVLSFNSVISTSLVLPVFPFAVVVFFHD 334
QY 305 KFOAERGVSLLSLMGFVSFYGFESKSK 334
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  335 RMNGVKIVAMLIANGFISVLFQHYLDGKK 364

RESULT 12
O49724 PRELIMINARY; PRT; 1128 AA.
AC O49724;
DT 01-JUN-1998 (T-EMBLrel. 06, Created)
DT 01-JUN-1998 (T-EMBLrel. 06, Last sequence update)
DT 01-OCT-2003 (T-EMBLrel. 25, Last annotation update)
DE Hypothetical protein.
GN T9A21.50 OR AT4G18200
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopses.
OX NCBI_TaxID=3702;
RN [1]
RP SEQUENCE FROM N.A.
RA Bevan M., Murphy G., Drost L., Hall C., Hudson S., Ridley P.,
RA Bancroft I., Mewes H.W., Mayer K., Schueller C.;
RL Submitted (FEB-1998) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RA Murphy G., Ridley P., Hudson S., Mewes H.W., Lemcke K., Mayer K.F.X.;
RA Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
RN [3]
RP SEQUENCE FROM N.A.
RA EU Arabidopsis sequencing project;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AL021713; CAA16792.1; -
DR EMBL; AL161548; CAB78822.1; -
DR PIR; T04922; T04922.
DR InterPro; IPR004853; DUF250.
DR Pfam; PF03151; DUF250; 3.
KW Hypothetical protein.
SQ SEQUENCE 1128 AA; 124621 MW; 981B88BFA0DE9E9B CRC64;

Query Match 29.7%; Score 533.5; DB 10; Length 1128;
Best Local Similarity 34.8%; Pred. No. 2e-33;
Matches 114; Conservative 72; Mismatches 133; Indels 9; Gaps 4;

QY 26 LYFTNGGKRIFWMSFLSTAGFPILIIPLIVSFLSRRGNRPNNAENKRTKFLMETPL 85
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  77 VYVENGKSTWMTLVQLIGFVPL---FLFRPFSQTK-NPKPTEADFRK----FSSFTIL 128
QY 86 FTASIVIGLTLGDNLYSYGLAYLPVSTSSLIIGTOLAFNALFALLVKQKFTPSINA 145
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  129 GSVYIVTGLLVANSYMSVGLLYLFPVSTFSLIASQLAFTAPFYSFLNSQKFTPTFVNS 188

```

```

QY 146 VLLTVGIGILALHSDGDKPAKESKKEYVVGFLMTVVAALLYAFILPLVELTYKKARQEI 205
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  189 LFLLTSSALLVVNTSENTAKVSRVKYVIGIITIGASAGIGLLLSLVQLILRKVLKKQ 248
QY 206 TPLVLEIQMVNCLAAATFCVGMIFVDFKVIAREAREPKIGSVFYVALIVITGIWQ 265
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  249 TFSVTVDLVAYOSLVASCVVILGLFASGEWKTLTSEMENYKL-GKVPYVNTLASIASISWQ 307
QY 266 GFPGLGAGIVFCASSLASGVLSVLLPVTVEFVAVVCFRKFQAEKGVSLLSLWGFVSF 325
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  308 VYTIGVVGILFESSVSFNSITAVGLPIPVVAVIVFHDKNASKIFSIILAIWGFISFV 367
QY 326 YGEFKSGKKVVDKQPPETELPILPVSD 353
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  368 YOHYLDKKLTSHTSFVGDPHLLPAEE 395

RESULT 13
Q8S170 PRELIMINARY; PRT; 372 AA.
AC Q8S170;
DT 01-JUN-2002 (T-EMBLrel. 21, Created)
DT 01-JUN-2002 (T-EMBLrel. 21, Last sequence update)
DT 01-JUN-2003 (T-EMBLrel. 24, Last annotation update)
DE Putative purine permease.
GN B1144G04.33 OR P0019E03.1.
OS Oryza sativa (japonica cultivar-group).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; Liliopsida; Poales; Poaceae;
OC Ehrhartoideae; Oryzeae; Oryza.
OX NCBI_TaxID=39947;
RN [1]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa nipponbare (GAJ) genomic DNA, chromosome 1, BAC
RT clone:B1144G04.";
RL Submitted (FEB-2001) to the EMBL/GenBank/DBJ databases.
RN [2]
RP SEQUENCE FROM N.A.
RC STRAIN=cv. Nipponbare;
RA Sasaki T., Matsumoto T., Yamamoto K.;
RT "Oryza sativa (japonica cultivar-group) genomic DNA, chromosome 1, PAC
RT clone:P0019E03.";
RL Submitted (NOV-2001) to the EMBL/GenBank/DBJ databases.
DR EMBL; AP003335; BAB90042.1; -
DR EMBL; AF004363; BAC01243.1; -
DR Gramene; Q8S170; -
DR InterPro; IPR004853; DUF250.
DR Pfam; PF03151; DUF250; 1.
SQ SEQUENCE 372 AA; 38794 MW; 6759730D536EC4AD CRC64;

Query Match 29.6%; Score 532; DB 10; Length 372;
Best Local Similarity 36.8%; Pred. No. 8.9e-34;
Matches 123; Conservative 67; Mismatches 122; Indels 22; Gaps 7;

QY 5 LIINCIIITGCGPLTRLYFTNGGKRIFWMSFLSTAGFPILIIPLIVSFLSRRGN 64
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  21 LUFANYAALGGSVASSLLSRYFFAHGHNWVATLVQSVGFPLVPVY-----AG 72
QY 65 RPNNAENKRTKFLMET-PLFIASIVIGLTLGDNLYSYGLAYLPVSTSSLIIGTOL 123
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  73 RSPS-----QRPFAWFTFRLLAACVIGLVGMVNNLLFSYSSVLPVSTSSLIISTOL 126
QY 124 AFNALFALLVKQKFTPFSSINAVLLTVGIGILALHS--DGDKPAKESKKEYVVGFLMTV 181
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  127 AFTLVLAIIVRHPLNFSNLNAVLLTSSVLIALRSSDSGEQPGGSRARYFVGFAVL 186
QY 182 VAALLYAFILPLVELTYKKARQEIFPVLVLEIQMVNCLAAATFCVGMIFVDFKVIARE 241
D  : : : : : : : : : : : : : : : : : : : : : : : : : : : : : :
Db  187 GAAGLFAAYLPVMELYVRKAVSG-GFRMAVEQVIMQNAATALAVAGLAAGGKW---EE 242
QY 242 AREFKIGGSVFYVALIVITGIITWQGFGLGAGIVFCASSLASGVLSVLLPVTVEFVAVV 301

```

```

Db      243 LARWDLSPAA-YWAVLAALVATWQACLMGTAGMYLTSSLHSGVCMATAVLTAHVIGVVV 301
QY      302 FREKFOAEKGYSLLSLWGFVSFYFGEKSKKV 335
Db      302 FRDPFGADKAVATVLCVWGFSSYLYGEYTTOKKV 335

RESULT 14
Q9ZUH3 PRELIMINARY; PRT; 315 AA.
AC Q9ZUH3;
DT 01-MAY-1999 (TrEMBLrel. 10, Created)
DT 01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT 01-OCT-2003 (TrEMBLrel. 25, Last annotation update)
DE AT2624220 protein.
GN AT2624220.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RX MEDLINE=20083487; PubMed=10617197;
RA Lin X., Kaul S., Rounsley S.D., Shea T.P., Benito M.-I., Town C.D.,
RA Fujii C.Y., Mason T.M., Bowman C.L., Barnstead M.E., Feldblyum T.V.,
RA Buell C.R., Ketchum K.A., Lee J.J., Ronning C.M., Koo H., Moffat K.S.,
RA Cronin L.A., Shen M., VanAken S.E., Umayam L., Tallon L.J., Gill J.E.,
RA Adams M.D., Carrera A.J., Cressy T.H., Goodman H.M., Somerville C.R.,
RA Copenhaver G.P., Preuss D., Nierman W.C., White O., Eisen J.A.,
RA Salzberg S.L., Fraser C.M., Venter J.C.;
RA "Sequence and analysis of chromosome 2 of the plant Arabidopsis
RT thaliana";
RL Nature 402:761-768(1999).
RN [2]
SEQUENCE FROM N.A.
RC STRAIN=cv. Columbia;
RA Lin X.;
RL Submitted (MAR-2000) to the EMBL/GenBank/DBJ databases.
DR EMBL; AC005967; A003383.1; -.
DR PIR; A84634; A84634.
DR InterPro; IPR004853; DUF250.
DR Pfam; PF03151; DUF250; 1.
SQ SEQUENCE 315 AA; 33847 MW; C76588FFABD4B3814 CRC64;

Query Match 29.6%; Score 531.5; DB 10; Length 315;
Best Local Similarity 36.8%; Pred. No. 8.2e-34;
Matches 114; Conservative 67; Mismatches 114; Indels 15; Gaps 2;

QY 22 LRLTYFTNGKRIWFMFSLTAGFPFIIPLLVSLFSLRRNRNPNNAENKRTKFLFM 81
Db 10 LLSRLYFSNGGKSWIISWAVAGWITCLILLPTVIF-----QKIKPTPL 55

QY 82 ETPLFTASIVIGLTDNLVYSYGLAYLPVSTSSLIIGTQLAFNALFAFLVKQKFTFP 141
Db 56 NTKLVLSVYVGLFSAADNLIMYAYAYLPASTSSLSLASSLAFSALFGLYIVKQPLNAS 115

QY 142 SINAVLLTVGIGLALHSDGDKPAKESKEYVVGFLMTVVAALLYAFILPVELTYKKA 201
Db 116 VINSIVITGAMAILIADSSDRYSYISNSQYFAGFFWDIMSGALHGLIFALSELVKL 175

QY 202 RQBITFPLVLEIQWMCCLAAATFFCVGMFIVGDFKVIAREAREFKIGGSFYALIVTG 261
Db 176 LGRSFHVALEQQWVSLTAFATFTTGMVVSNDFOGMSHEAKSF-K-GGESLYTQVILWNSA 234

QY 262 IINQGFPLGAIGVFCASSLASGVLSVLLPVTVEAVVCFREKFOAEKGVSLLSLWGF 321
Db 235 VTFQGLVGTAVLFLASTVWAGVLNAVVRPITTSVAAILMHDFMSGFKILSLVLFWNGF 294

QY 322 VSYFYGEFKS 331

```

```

Db      295 SSVYIGSSSS 304

RESULT 15
Q8RY74 PRELIMINARY; PRT; 377 AA.
AC Q8RY74;
DT 01-JUN-2002 (TrEMBLrel. 21, Created)
DT 01-JUN-2002 (TrEMBLrel. 21, Last sequence update)
DT 01-OCT-2002 (TrEMBLrel. 22, Last annotation update)
DE Hypothetical protein.
GN A74GI8200.
OS Arabidopsis thaliana (Mouse-ear cress).
OC Eukaryota; Viridiplantae; Streptophyta; Embryophyta; Tracheophyta;
OC Spermatophyta; Magnoliophyta; eudicotyledons; core eudicots; rosids;
OC eurosids II; Brassicales; Brassicaceae; Arabidopsids.
OX NCBI_TaxID=3702;
RN [1]
SEQUENCE FROM N.A.
RA Yamada K., Liu S.X., Sakano H., Pham P.K., Banh J., Chung M.K.,
RA Goldsmith A.D., Lee J.M., Quach H.L., Toriumi M., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Miranda M., Narusaka M., Nguyen M., Palm C.J., Sakurai T., Sacou M.,
RA Seki M., Shinn P., Southwick A., Shinozaki K., Davis R.W., Ecker J.R.,
RA Theologis A.;
RA "Arabidopsis Full Length cDNA Clones";
RT Submitted (JAN-2002) to the EMBL/GenBank/DBJ databases.
RN [2]
SEQUENCE FROM N.A.
RA Yamada K., Banh J., Chan M.M., Chang C.H., Chang E., Dale J.M.,
RA Deng J.M., Goldsmith A.D., Lee J.M., Onodera C.S., Quach H.L.,
RA Tang C., Toriumi M., Wu H.C., Yamamura Y., Yu G., Bowser L.,
RA Carninci P., Chen H., Cheuk R., Hayashizaki Y., Ishida J., Jones T.,
RA Kamiya A., Karlin-Neumann G., Kawai J., Kim C., Lam B., Lin J.,
RA Meyers M.C., Miranda M., Narusaka M., Nguyen M., Palm C.J.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA Sakurai T., Satou M., Seki M., Shinn P., Southwick A., Shinozaki K.,
RA "Arabidopsis Open Reading Frame (ORF) Clones";
RL Submitted (APR-2002) to the EMBL/GenBank/DBJ databases.
DR EMBL; AY074546; AAL69512.1; -.
DR EMBL; AY096558; AAM20208.1; -.
DR InterPro; IPR004853; DUF250.
DR Pfam; PF03151; DUF250; 1.
KW Hypothetical protein.
SQ SEQUENCE 377 AA; 41786 MW; B77A09A4289859AB CRC64;

Query Match 29.3%; Score 526.5; DB 10; Length 377;
Best Local Similarity 35.2%; Pred. No. 2.4e-33;
Matches 118; Conservative 70; Mismatches 124; Indels 23; Gaps 7;

QY 21 PLLT---RLYFTNGKRIWFMFSLTAGFPFIIPLLVSLFSLRRNRNPNNAENKRTK 77
Db 54 PLATVGLRLYENGKSTYVVTLLQLLGFPVL---ILFRFSRIR-----QPKSTD 101

QY 78 LFLMETPLF--IASIVI--GLLTGLDNLVYSYGLAYLPVSTSSLIIGTQLAFNALFAFL 133
Db 102 TNFSQSPSFTTLASVYLCTGLVSAVAYLSAVGLLYLPVSTFSLILASQLAFTAFSFL 161

QY 134 VKQFTPSINAVLLTVGIGLALHSDGDKPAKESKEYVVGFLMTVVAALLYAFILPL 193
Db 162 NSQKFTPLVNSLFLTLVSSALLVNTDSENTVNSRVQYVIGFTICTGASAGLGLVSL 221

QY 194 VELTVKKAQETTFPLVLEIQWMCCLAAATFFCVGMFIVGDFKVIAREAREFKIGGSV 253
Db 222 IQLLFRKFTKHTTSVALLDLANYQSLVATCVVLIGLFASGEWRTLPSERNYKL-GKVS 280

QY 254 YALIVITGIINQGFPLGAIGVFCASSLASGVLSVLLPVTVEAVVCFREKFOAEKGV 313
Db 281 ILTASAAIFWQVYVGVGVLFIFESSVSFNSITAVGLPIVFWVAVIVFHDKMDASKIFS 340

QY 314 LLLSLWGFVSFYFGEKSKKV---VDKPPQPETE 345

```

Db 341 IILAINGFLSFVYQHVLDEKKLKTCTQKPVREBTO 375

Search completed: September 23, 2004, 23:49:45  
Job time : 77.4181 secs

**THIS PAGE BLANK (USE TO)**